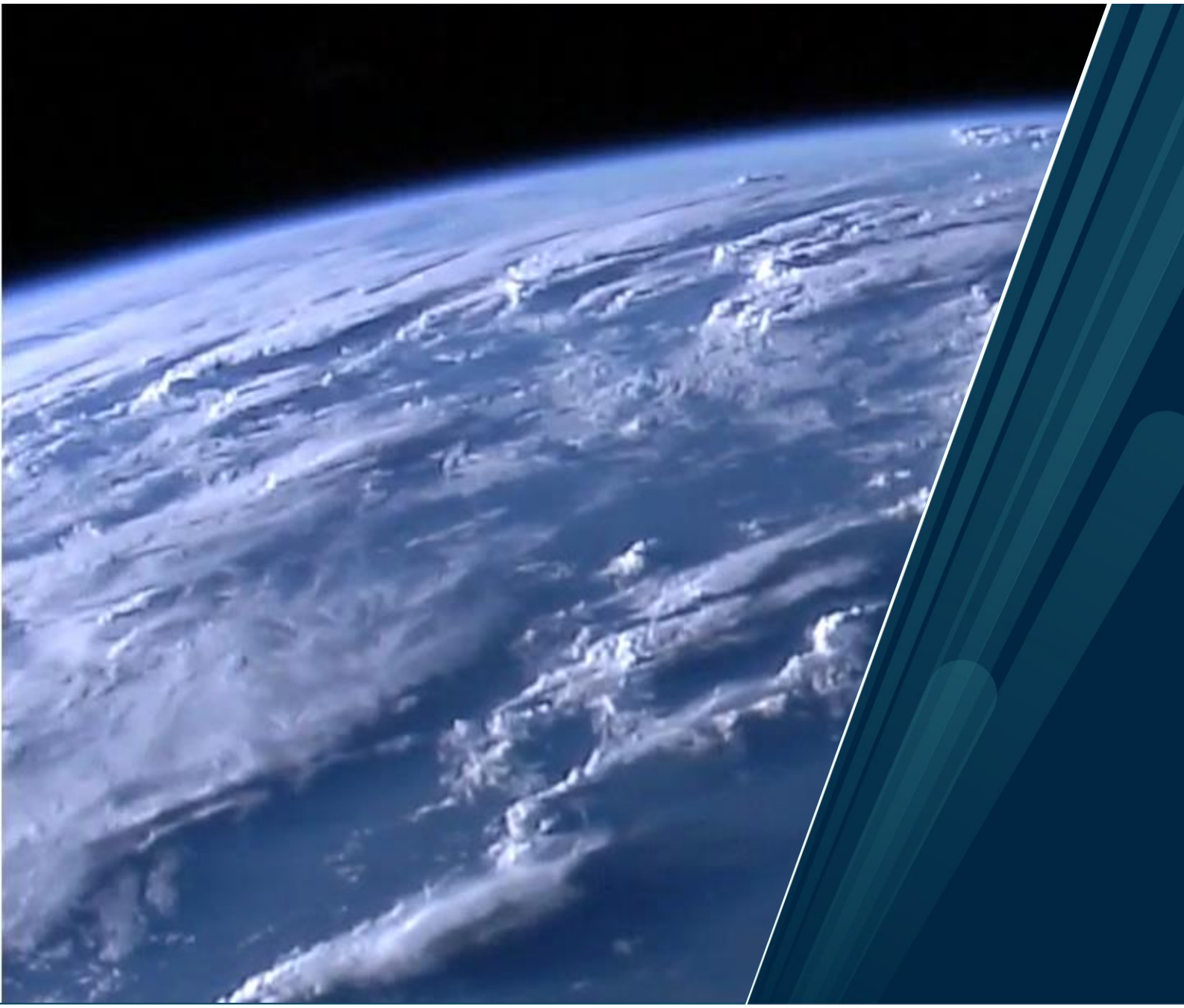
UiT The Arctic University of Norway

Department of Technology and Safety

Simulation of Meteorological and Oceanographic Parameters: An Application in Spray Icing Prediction

Abolfazl Shojaei Barjouei

Master's thesis in Technology and Safety in the High North…TEK - 3901…July 2020



Abstract

Sea spray icing is considered as a major environmental challenge in the Arctic Ocean, which poses a critical risk not only to the vessels and industrial operations but also to human safety. Although some studies have been carried out to estimate spray icing rate (e.g., RIGICE04 and ICEMOD models), such models suffer from some unrealistic modeling assumptions and limited verification. Moreover, limited researches have been conducted on the prediction of icing rates in the long-term, as well as climatological information on spray icing for long-term risk-based decisions in the Arctic offshore industrial applications. In this study, simulation of meteorological conditions to improve prediction of sea spray icing for offshore industrial applications in the Arctic region is purposed. The applications of Bayesian inference as well as Monte Carlo methods comprised of Sequential Importance Sampling (SIS) and Markov Chain Monte Carlo (MCMC) in the prediction of meteorological and oceanographic parameters to improve the estimation of sea spray icing in the Arctic region is purposed. Reanalysis data from NOrwegian ReAnalysis 10km (NORA10) during 33 years are applied to evaluate the performance of the models. Consequently, using the 32-year data, the parameters are predicted and compared for the last one-year on a daily basis. The predicted parameters are considered as input for the newly introduced icing model namely Marine-Icing Model for the Norwegian COast Guard (MINCOG) and the results are evaluated and discussed. Apart from the prediction of sea spray icing, the applied prediction and simulation techniques can play useful roles in industrial application, especially, when new data and information are collected using which the meteorological and atmospheric conditions are predicted for future junctures. This provides the decision-maker with valuable information for planning offshore activities in the future (e.g., offshore fleet optimization). Accordingly, sea voyages with relatively lower risks can be selected based on the predicted parameters and icing rates.

Keywords: Marine icing, Arctic offshore, Meteorology, Oceanography, Simulation, Bayesian approach, Sequential importance sampling, Markov chain Monte Carlo.

Table of Contents

List of Tables

List of Figures

Foreword

This thesis has been carried out as the final part of a two-year international master program in "Technology and Safety in the High North" at the Department of Technology and Safety at UiT The Arctic University of Norway.

The study was initiated after a former specialization project entitled "Simulating the Icing Evens and Rates in Future Junctures in the Arctic: A Bayesian Approach". There were abundant opportunities for further research and development in expanding the Bayesian inference and examining other simulation approaches in the study subject. Consequently, and after exploring among tens of recently published articles and books, the modern simulation techniques that have been derived from the Monte Carlo simulation method seemed to be more relevant, which along with Bayesian inference are considered for detailed investigation in the current study.

To make this research simple and straightforward so that researchers, students in science and engineering as well as engineers in industries be able to apply or further develop the techniques, it is tried to cover the required information as much as possible. However, basic information from the course of general statistics and a programming language are initially required.

Acknowledgment

Hereby, I would like to express my gratitude to my supervisor, Assistant Professor Masoud Naseri, for wise ideas, useful tips in programming, technical feedbacks, and contributions during this research.

My sincere thanks also go to my co-supervisor, Professor Javad Barabady, for the inspiring guidance, ongoing support, and useful comments not only during this research but also over the entire study program.

Moreover, I would like to acknowledge the Norwegian Deepwater Programme for the use of NORA10 hindcast data.

Abbreviations

ACO	Ant Colony Optimization
CDF	Cumulative Distribution Function
CFD	cumulative frequency distribution
CV	Coefficient of Variation
ECDF	Empirical Cumulative Distribution Function
EDA	Exploratory Data Analysis
HPD	Highest Posterior Density
i.i.d.	Independent and identically distributed random variables
IS	Importance Sampling
MCDM	Multi-Criteria Decision-Making
MCMC	Markov Chain Monte Carlo
MCS	Monte Carlo Simulation
MET	Norwegian Meteorological Institute
MINCOG	Marine-Icing Model for the Norwegian COast Guard
MLE	Maximum Likelihood Estimator
NORA10	NOrwegian ReAnalysis 10km
PDF	Probability Density Function
RAMS	Reliability, Availability, Maintainability, and Safety
SAR	weather and ocean conditions aid the authorities in the Search and Rescue
SIR	Sampling Importance Resampling
SIS	Sequential Importance Sampling
SMC	Sequential Monte Carlo

1 Introduction

Spray icing is considered a major environmental challenge and critical risk element for offshore activities in the Arctic waters. Icing may impact offshore operations, reduce safety, operational tempo and productivity, cause malfunction of the operational and communication equipment, slippery handrails, ladders or decks, unusable fire and rescue equipment, and the blocking of air vents (Ryerson, 2011; Dehghani-Sanij, Dehghani, Naterer, & Muzychka, 2017). The icing on vessels may also lead to severe accidents and capsizing (Heinrich, 1950; Chatterton & Cook, 2008). Two main sources of icing are sea spray due to collision of ship and waves, as shown in Figure 1, and atmospheric icing caused by fog, Arctic sea smoke, high-velocity wind, and rain/drizzle or snow, as shown in Figure 2 (Rashid, Khawaja, & Edvardsen, 2016).

Wide varieties of techniques and technologies are available to enhance icing safety and protection such as chemicals, coatings, heat, and high-velocity fluids, air, water, and steam (Rashid, Khawaja, & Edvardsen, 2016). However, forecasting the amount and frequency of ice formation aids the selection of safety-enhancing strategy and ice protection technologies. Forecasting, also, can aid in tactical preparation before an icing event (Ryerson, 2011).

Figure 1. Icing due to the collision of ship and waves (Toomey, Lloyd, House, & Dickins, 2010)

Figure 2. Atmospheric icing (Overland, 2000)

Nevertheless, forecasting icing events and rate is a complicated task due to the chaotic nature of icing and its correlation with a large number of parameters. Extensive works have been conducted on historical icing data to predict icing events (Samuelsen, Edvardsen, & Graversen, 2017). Accordingly, the data have been examined from different perspectives such as the influence of meteorological parameters on icing rate from the statistical point of view (Mertins, 1968), and introducing sea spray algorithms based on the collision of ship and waves considering the environmental data as input parameters (Stallabrass, 1980; Samuelsen, Edvardsen, & Graversen, 2017). The history and development of sea spray icing predictive models are reviewed by Dehghani-Sanij et al. (2017) and Sultana et al. (2018). Accordingly, to estimate the icing rate on a vessel, ICEMOD (Horjen, 1960; Horjen, 2013) and RIGICE04 (Forest, Lozowski, & Gagnon, 2005) are commonly used prediction models. However, the newly developed model, MINCOG, provides higher verification scores than previously applied vessel-icing models. The MINCOG model is developed based on the modeling of sea spray from wave-vessel interaction, which is considered as the main water source in vessel-icing

events (Samuelsen, 2018). In the MINCOG model, six meteorological and oceanographic parameters comprised of wave height, wind speed, temperature, relative humidity, atmospheric pressure, and wave period are to be predicted as inputs. Apparently, lack of accuracy in the input parameters leads to uneven results. Therefore, enhancing the MINCOG results through improving the prediction of required input parameters is considered in this study. Meanwhile, the predicted meteorological and oceanographic parameters can also be utilized in any other industrial application or research study.

As a matter of climate change, direct use of old data does not lead to accurate predictions of the future. One approach to deal with less reliable historical data might be ignoring the older values and only using more recent data. Elsner and Bossak (2001) proposed another alternative applying the Bayesian approach, in which different qualities were considered for earlier and more recent data. They applied the approach for the prediction of U.S. landfalling hurricanes. Accordingly, a prior distribution was estimated using earlier data from 1851 to 1899. Then, the remaining data from 1900 to 2000 were used to revise the prior distribution. In this regard, Bayesian inference is considered a strong approach to deal with the uncertainties in meteorological and oceanographic parameters. Wang et al. (2019) applied the Bayesian approach in a model to estimate the uncertainties associated with weather and climate projections (e.g., 2-m temperature, surface radiation fluxes, or wind speed). Zhao et al. (2017) used a Bayesian statistical technique to improve the accuracy of temperature fields. Rainfall prediction utilizing a Bayesian approach was considered by Nikam and Meshram (2013). Wikle et al. (2013) reviewed a heterogeneous mix of studies demonstrating Bayesian hierarchical model applications in ocean physics, air-sea interaction, ocean forecasting, and ocean ecosystem models. Cornejo-Bueno et al. (2018) applied a Bayesian approach to obtain the optimal parameters of a prediction system for problems related to ocean wave features prediction. Alternatively, Monte Carlo simulation can also be used to simulate icing rate as well as meteorological and oceanographic parameters (Naseri & Samuelsen, 2019; Ali, Deo, Downs, and Maraseni, 2020). Thus, in this study application of Bayesian inference as well as simulation approaches comprised of SIS and MCMC for long-term prediction of the meteorological and oceanographic parameters is considered. The predicted parameters will be further used as inputs for the MINCOG model for predicting the rate of sea spray icing.

1.1 Applications

The main application of this study is in risk assessment and identifying dangerous situations and consequently risk-based decision-making. The outcomes of this study provide decision-makers with information for both short- and long-term planning. For instance, the results aid in assigning annual budgets, hiring crew for future purposes (e.g., overhaul in the next six months), and facilitating the platform/vessel to mitigate the risk of encountering critical situations in the future. The results also assist to choose between alternative sea routes for a safer voyage of offshore supply vessels, fishing vessels, and cruise ships (Naseri & Samuelsen, 2019).

Ice accretion on vessels can significantly increase the load and consequently, the fuel consumption will be increased. The fuel consumption is also affected by the energy that is required for deicing purposes. Therefore, neglecting the possible ice accretion and preparing the required fuel, the vessels may be encountered with the threat of being out of fuel, which particularly in the Arctic waters can cause catastrophic consequences. Furthermore, considering icing rates and fuel consumption in long-term planning results in better decisions and ultimately increase profits. Ice accumulation also influences the operability of vessels as well as offshore production structures and facilities by increasing power losses, failure rate, and frequency of need for inspection and repair. Additionally, it reduces the useful lifetime of the equipment and imposes safety hazards (Barabadi, Garmabaki, & Zaki, 2016). Hence, accurate prediction of icing rate aids to decide mitigating measures.

Maintainability of the equipment is another challenge of vessels and industrial activities in the Arctic. In this regard, long-term prediction of icing rate and other environmental conditions such as wave height, temperature, and wind speed is essential for proper maintenance planning and scheduling.

The reliability of equipment and facilities on-board are also highly affected by ice and environmental conditions, especially low temperatures and high-velocity winds, in the Arctic. Wind chill effect removes heat from equipment and imposes more fatigue on the equipment. Therefore, appropriate simulation of the future environmental condition can enhance the design of equipment for the future.

Preparing plans for industrial activities and sea voyages in the Arctic without considering the predictions of the condition would not result in proper outcomes. In the planning phase, it is necessary to consider the conditions of the date of operations and additionally the availability of the equipment in that condition. Otherwise, long delays, extra expenses, serious injuries, and perhaps fatalities are expected.

Due to global warming and ice melting in the Arctic, the marine traffic for both industrial and leisure purposes has largely been increasing, particularly in the area around Svalbard. In such a situation, accurate long-term predictions of the weather and ocean conditions aid the authorities in the Search and Rescue (SAR) sector for better planning and preparation for future junctures and properly reviewing and modifying the rules and regulations. For instance, the SAR exercises in Spitzbergen in 2016, 2017 and 2018 can be referred that were carried out by the Norwegian coast guard in cooperation with a large number of Norwegian and international entities, governmental, official and private (SARex1, 2016; SARex2, 2017; SARex3, 2018).

1.2 Aims

Firstly, enhancing the long-term prediction of sea spray icing resulted by the MINCOG model through improving the prediction of required input parameters including wave height, wind speed, temperature, relative humidity, atmospheric pressure, and wave period is purposed.

Secondly, investigating the applications and performances of statistical and simulation approaches such as Bayesian inference, Monte Carlo simulation, importance sampling, sequential Monte Carlo, and resampling dealing with forecasting meteorological and oceanographic parameters is considered.

Thirdly, the predicted icing rates and meteorological and oceanographic parameters provide useful information that is essential for developing frameworks to enhance Reliability, Availability, Maintainability, and Safety (RAMS).

Finally yet importantly, considering the variety of approximation techniques that are summarized in the methodology, this study provides a useful resource for students of engineering and sciences as well as researchers who are interested in data analysis, estimation, simulation, and forecasting.

1.3 Scope

The scope of this study is the prediction of meteorological and oceanographic parameters for a better forecast of spray icing encountered with offshore activities in the Arctic. The purposed area of study is the sea area between Northern Norway and Svalbard archipelago, bounded to the latitudes 69°N to 78°N and longitudes 8°E to 36°E. However, the parameter prediction and simulation methodologies can be applied for onshore activities on any other location.

1.4 Limitations

In order to validate the performances of the purposed frameworks, 33 years of reanalysis data from NOrwegian ReAnalysis 10km (NORA10) including meteorological and oceanographic parameters such as wave height, wind speed, temperature, relative humidity, atmospheric pressure, and wave period from 1 January 1980 to 31 December 2012 is utilized. However, the frameworks are applicable for updated datasets and other areas of interest.

1.5 Structure

The structure of the study is organized as follows. Section 2 is devoted to the methodology in which exploratory data analysis, Bayesian inference as well as SIS and MCMC techniques are presented in detail. In Section 3, the available dataset is explained, and considering some assumptions, the models are modified and applied to the dataset to simulate the meteorological and oceanographic parameters in the study area. Eventually, the predicted data are used as input parameters in the MINCOG model to forecast the icing rate and the results are discussed. Finally, conclusions and possible future studies are mentioned in Section 4.

2 Methodology

When attempting to anticipate future patterns of meteorology and oceanography under varying options, inevitably a model of reality is required, which can never fit reality in all details. The model is based on the available information on the behavior of different parameters of the system and their interactions. Accordingly, how the parameters of the system move among their possible states and, eventually, how the system behaves can be determined (Zio, 2013).

Making sense of a new dataset is one of the applications of the statistical concept in meteorology and climatology. Torrents of numerical data are being produced by meteorological observation

systems and computer models, which make it a critical task to get a feel about batches of numbers and to extract insight into underlying generating processes. In this view, statistical inference draws conclusions about the characteristics of a "population" based on a limited data sample. In other words, inferential methods aim to extract the generating process of the data sample. In this context, two main approaches are considered by statisticians including Frequentist and Bayesian inferences. Frequentist inference deals with the distribution that can well describe the data at hand. Meanwhile, this distribution can be parametric or nonparametric. In Bayesian inference, however, a parametric distribution is assumed to characterize the nature of the data-generating process and the parameter(s) of the distribution are the subject of uncertainty. Accordingly, prior information about the parameter(s) of interest is quantified by a probability distribution, which may or may not be of a familiar parametric form. This prior information is then modified by combining with the information provided by the data sample, in an optimal way (Little, 2006; Wilks, 2011). Consequently, Bayesian inference provides a proper understanding of the stochastic nature of the parameter(s), although its calculations are more complex than Frequentist inference. Therefore, and given the credibility and ease in the model development procedure, the Bayesian inference has been recognized as a promising analysis technique to tackle the events with chaotic nature (Park, Ju, & Kim, 2020).

Alternatively, the Monte Carlo Simulation (MCS) method is the other powerful modeling tool dealing with complex and chaotic events to achieve a closer adherence to reality. MCS is generally defined as a methodology to estimate the solution of mathematical problems using random numbers. Taking advantage of the present powerful computers, the MCS method is still becoming more and more practicable in various fields such as simulation of random walks in a naturally stochastic environment, and approximating the solution of equations, both differential and integral. Standard MCS was also enhanced by drawing conditional samples by means of a Markov chain so-called MCMC (Robert & Casella, 2012; Zio, 2013). MCMC is particularly beneficial in estimating Bayesian posterior, which determines the probability or density value of an event given relevant evidence (Tierney, 1994). In this view, MCMC seems to be relevant to the purpose of this study in which simulating the future behavior of meteorological and oceanographic parameters given past evidence is considered.

2.1 Exploratory Data Analysis

Getting a feel for a new batch of numbers and extract insight about the processes underlying their generation is broadly known as *Exploratory Data Analysis* (EDA), which taking advantage of graphical methods aids in the comprehension of the large sets of data that may confront an analyst (Tukey, 1977). Graphical methods effectively aid to compress and summarize data. Consequently, they can be portrayed in little space, which makes the comprehension of the large batches of numbers straightforward. Moreover, graphically oriented computer packages have made the use of these methods fast and easy (Wilks, 2011).

Stringent assumptions about the nature of the data such that data will follow the familiar bell-shaped curve of the Gaussian distribution are common in the classical techniques of statistics while if the assumptions are not provided by the data might lead to quite misleading results. However, simplifying assumptions allows deriving elegant analytic results that are mathematically powerful but relatively simple. Accordingly, two important aspects that reduce the sensitivity of the nature of a dataset to assumptions are robustness and resistance. Robustness does not necessarily lead to optimality in any particular circumstance while leading to a reasonable performance in most circumstances. For instance, although the sample average is the best features of the center of dataset assuming the data are following a Gaussian distribution, if the data do not satisfy the assumption, the sample average will lead to a misleading of centralization feature. However, a robust method is not generally dependent on particular assumptions regarding the overall nature of the data. Resistant methods are those their results are not dependent on the small number of outliers. In other words, the results of a resistant method hardly fluctuate even in case of drastic changes in the data values (Wilks, 2011).

Location, spread, and symmetry are three simple robust and resistant summary measures that can be applied without plotting or computer graphic capabilities. The location represents the central tendency of the data values while spread refers to the degree of variation around the center. Additionally, the balance of data distribution around the center is described by the spread. However, the classical measures of these three measures including the sample mean,

sample variance, and sample coefficient of skewness, respectively are neither robust nor resistant (Wilks, 2011).

2.1.1 Density Estimation

In EDA, estimating the distribution of the data is a useful presentational tool by providing a very effective means of compressing and summarizing data, portraying much in little space, and exposing features of the data such as central tendency, spread, symmetry, and percentiles (Wilks, 2011). An estimate of the density function is also useful for decision-making, classification, and summarizing Bayesian posteriors. Furthermore, density estimation can be considered as a tool in other computational methods, such as some simulation algorithms and Markov chain Monte Carlo approaches (Silverman, 1986; Scott, 2015).

The parametric solution to estimate a density involves the risk of relying on an incorrect model that can lead to serious inferential errors, regardless of the estimation strategy used to generate the parameter (e.g., maximum likelihood, Bayesian, or method-of-moments). In this regard, estimating nonparametric densities is an interesting concept in statistics due to the fact that for most real-world problems a proper parametric form of density is either unknown or does not exist (Givens & Hoeting, 2013).

Histogram is a common nonparametric density estimator, which is a piecewise constant density estimator. The range of the data is divided into class intervals so-called bins, which widths are defined by the class limits, and the heights depend on the number of values falls in each bin. Consequently, attributes of the data distribution as location, spread, and symmetry are revealed. However, the choice of bin width is the main issue to construct a proper histogram. Too wide intervals result in too smooth histogram in which important details of the data might be masked. Whilst, too narrow intervals result in too rough plot that is difficult to interpret. Moreover, it is required to round each data value to the center of the bin into which it falls (Wilks, 2011).

2.1.2 Kernel Density Estimation

An extension of the histogram that does not require arbitrary rounding to bin centers, and that provides a smooth result, is the *kernel density estimate*. Indeed, *kernel density estimate* is a nonparametric alternative to fit the common parametric probability density functions.

The kernel height for a given value, x_0, corresponding to the data values $x_i; i = 1, \dots, n$, is calculated as below (Wilks, 2011):

$$\hat{f}(x_0) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x_0 - x_i}{h}\right) \quad (1)$$

where K is a *smoothing kernel function* and h is the smoothing parameters, also known as the bandwidth. *Smoothing kernel function* is a nonnegative function with unit area, that is, $\int K(t)dt = 1$, so each is a proper Probability Density Function (PDF). Moreover, it is centered at zero. Meanwhile, n is the number of data values that are close enough to the point x_0 (i.e. the distances to x_0 is closer than h) to result in non-zero kernel height. Some of the commonly used *smoothing kernel functions* for continuous data are shown in Table 1 and plotted in Figure 3. Meanwhile, some functions that are appropriate to discrete data are presented by Rajagopalan et al. (1997).

Table 1. Commonly used smoothing kernel functions (Wilks, 2011)

Name	$K(t)$	Support
Quartic (Biweight)	$(15/16)(1-t^2)^2$	$-1 < t < 1$
Triangular	$1 - \lvert t \rvert$	$-1 < t < 1$
Quadratic (Epanechnikov)	$(3/4)(1-t^2)$	$-1 < t < 1$
Gaussian	$2\pi^{-1/2}\exp[-t^2/2]$	$-\infty < t < \infty$

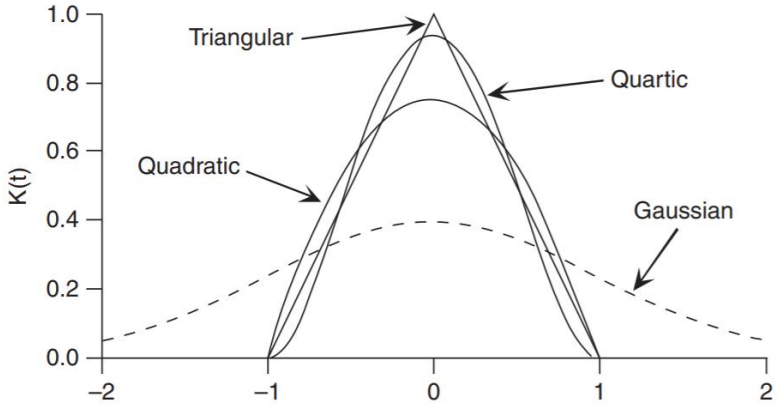


Figure 3. Commonly used smoothing kernels functions (Wilks, 2011)

2.1.3 Empirical Cumulative Distribution Function

The *cumulative frequency distribution*, also known as the *Empirical Cumulative Distribution Function* (ECDF), is a two-dimensional plot associated with the histogram. The vertical axis in the cumulative frequency distribution shows cumulative probability estimates related to the data values on the horizontal axis. Indeed, the plot shows the estimation of the relative frequency for the probability that a random future datum will not exceed the corresponding value on the horizontal axis. Hence, the cumulative frequency distribution can be interpreted as the integral of a histogram with an arbitrarily narrow bin width. Meanwhile, ECDF can be smoothed obtained integrating the result of a kernel smoothing, just like the kernel density smoothing that applied to histograms.

The vertical axes in Figure 4 indicates ECDF, $p(x)$ that is expressed as

$$p(x) \approx P(X = x) \tag{2}$$

where $P(e)$ represents the probability that the event e happens.

To construct a cumulative frequency distribution, $p(x)$ must be estimated using the ranks, i, of the order statistics, $x_{(i)}$. These estimates are known as *plotting positions* that are historically used in graphically comparing the empirical distributions with candidate parametric functions (Harter, 1984). A substantial literature has been devoted to calculate plotting positions and thus to estimate cumulative probabilities from datasets, which are mostly a particular case of the formula

$$p(x_i) = \frac{i - a}{n + 1 - 2a}, 0 \leq a \leq 1 \tag{3}$$

Accordingly, different plotting position estimators are resulted by different values for the constant a, as some of them are shown in Table 2. It should be mentioned that the names of the functions in this table are taken from authors who proposed the various estimators, not from particular probability distributions. Thus, using the Weibull plotting position estimator, the ECDF for daily average temperature in coordinate (69.3°N, 8.6°E) on the 1st of January during

the years 1980 to 2012 is shown Figure 4, which is a step function with probability jumps occurring at the data values.

Table 2. Common plotting position estimators (see Equation 3) for cumulative probabilities (Wilks, 2011)

Name	Formula	a	Interpretation
Weibull	$i/(n+1)$	0	Mean of sampling distribution
Benard & Bos-Levenbach	$(i-0.3)/(n+0.4)$	0.3	Approximate median of sampling distribution
Tukey	$(i-1/3)/(n+1/3)$	$1/3$	Approximate median of sampling distribution
Gumbel	$(i-1)/(n-1)$	1	Mode of sampling distribution
Hazen	$(i-1/2)/n$	$1/2$	Midpoints of n equal intervals on $[0,1]$
Cunnane	$(i-2/5)/(n+1/5)$	$2/5$	Subjective choice commonly used in hydrology

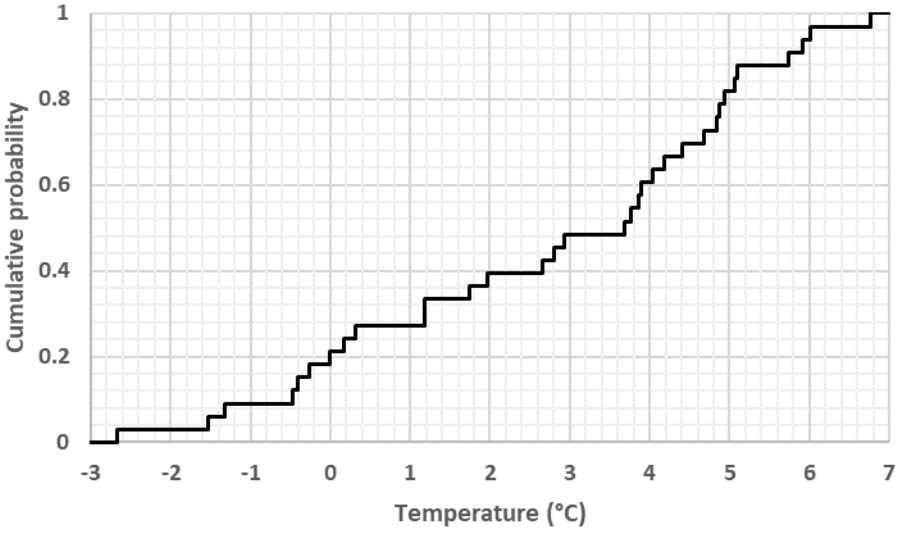


Figure 4. ECDF using Weibull plotting position estimator for daily average temperature in coordinate (69.3°N, 8.6°E) in 1st of January from 1980 to 2012

2.2 Bayesian Inference

The Bayesian inference is a parametric view of probability in which the parameters of probability distributions are the subject of inference. A parametric distribution quantitatively characterizes the dependency of the nature of the data-generating process on the parameter(s) about which inferences are being drawn. For instance, if the data have achieved through *N* identical and independent Bernoulli trials, the binomial distribution can be considered as the data-generating model and the binomial parameter, *p*, is the target of statistical inference, which can fully describe the nature of the data-generating process (Pole, West, and Harrison, 1994; Walshaw, 2000; Elsner and Bossak, 2001).

2.2.1 Bayes' Theorem

Regardless the variable of interest is discrete or continuous the parameter that is subject of inference is generally continuous and can be presented as probability density function. Accordingly, Bayes' Theorem for continuous probability models can be represented as follows (Wilks, 2011):

$$f(\theta|x) = \frac{f(x|\theta)f(\theta)}{f(x)} = \frac{f(x|\theta)f(\theta)}{\int_{\theta} f(x|\theta)f(\theta)\,d\theta} \tag{4}$$

where, θ is the distribution parameter (e.g., p in the binomial distribution or λ in a Poisson distribution), and *x* is the data in hand. Subjective belief about the parameter θ is described by the prior distribution $f(\theta)$ which is generally a PDF when θ is a continuous parameter. However, different forms of $f(\theta)$ may be chosen by different analysts (Wilks, 2011). Furthermore, the likelihood, $f(x|\theta)$, represents the general nature of the data-generating process, which will be influenced by different values of θ on it. It is worth mentioning, the likelihood is in fact a function of the parameter θ based on fixed values of the data *x* rather than a function of data *x* based on fixed-parameter θ. In other words, $f(x|\theta)$ expresses the relative plausibility of the data as a function of possible values of θ. Consequently, the posterior distribution, $f(\theta|x)$, is resulted by updating the prior distribution $f(\theta)$ considering the provided information by the likelihood (Wilks, 2011).

2.2.2 Posterior Distribution

The basis of the statistical inference in the Bayesian structure is provided by the posterior distribution, $f(\theta|x)$. In case the posterior distribution is not of a common parametric form, a point estimation of the parameter θ such as central tendency measures including the mean, median, or mode of the posterior distribution might be of interest. Particularly, the posterior mode is an attractive point estimator due to its relationship to the Maximum Likelihood Estimator (MLE) of θ (Wilks, 2011). The influence of the prior distribution on the posterior distribution decreases as the amount of data rises which makes the posterior distribution almost proportional only to the likelihood. Then, the posterior mode is approximately the same as the MLE of θ. However, summarizing the posterior distribution via probabilities seems more informative rather than a point estimation of central tendency. Therefore, a central credible interval is usually considered, which extends a range for θ analogous to the middle portion of the posterior distribution. An alternative for this credible interval is the Highest Posterior Density (HPD) interval which is defined with respect to the largest possible related values of the posterior distribution. The HPD interval can also be considered as a probabilistic extension of the posterior mode. The HPD interval coincides with the simple central credible interval for a symmetric posterior distribution while for a skewed distribution it will be shifted and compressed (Wilks, 2011).

2.2.3 Prior Distribution

The analyst's degree of belief or uncertainty regarding possible values of θ before new data arrives is quantitatively characterizes through prior distribution $f(\theta)$. This latter point makes the Bayesian inference controversial especially when the available data are relatively few since different priors, depending on the analyst's judgment, cause quite different posteriors. However, as the number of data increases, this dependency decreases and similar inferences will be derived from reasonable priors. Another aspect of the prior distribution is that it is not necessarily required to be of familiar parametric forms, although adopting a known parametric form is both conceptually and mathematically convenient and may considerably simplify the subsequent calculations (Wilks, 2011).

2.2.4 Predictive Distribution

Gaining insight about unobserved data values (x^+) in the future by quantifying the uncertainty of the parameter θ is the ultimate goal of Bayesian inference. To this aim, a probability density function namely predictive distribution is derived by combining the parametric data-generating process and the posterior distribution for θ, which is given by Equation 5 (Wilks, 2011),

$$f(x^+) = \int_\theta f(x^+|\theta) f(\theta|x)\, d\theta \tag{5}$$

where x^+ represents the unobserved data in the future and x denotes the data in hand which has already been used to derive the posterior distribution $f(\theta|x)$. It should be noted that $f(x|\theta)$ is the PDF for the data given a particular value of θ, not the likelihood for θ given a fixed data sample x, although the two have the same notation. The posterior PDF $f(\theta|x)$ quantifies uncertainty regarding θ based on the most recent probability updates. Equation 5 is indeed a weighted average of the PDFs $f(x^+|\theta)$ for all possible values of θ, where the posterior distribution provides the weighs (Wilks, 2011).

2.2.5 Gaussian Data-Generating Process

Applying Bayesian inference in which the distribution of parameters of the generating process (i.e., $f(\mu)$ and $f(\sigma^2)$) are unknown is quite complicated and out of the scope of the current work. The treatment to cope with the situation is available in (Epstein, 1985; Lee, 1997). The other alternative case is the assumption of the known variance of the data generating process for inferences about a Gaussian μ. Additionally, considering the later assumption in the situation where the conjugate prior and posterior distributions are Gaussian is computationally convenient, although it is confusing in notation due to four sets of means and variances. Moreover, when the posterior is Gaussian, the predictive distribution will also be Gaussian (Wilks, 2011). The sets of means and variances are as follows:

- μ: mean of the data-generating process
- σ_*^2: known variance of the data-generating process
- (μ_h, σ_h^2): hyperparameters of the prior Gaussian distribution
- $\bar{x}$: sample mean

- $(\mu_h', \sigma_h^{2\prime})$: hyperparameters of the posterior Gaussian distribution
- (μ_+, σ_+^2): parameters of the Gaussian predictive distribution

Accordingly, the prior distribution is proportional to Equation 6 (Wilks, 2011).

$$f(\mu) \propto \frac{1}{\sigma_h} exp\left[-\frac{(\mu-\mu_h)^2}{2\sigma_h^2}\right] \tag{6}$$

Moreover, given a sample of *n* independent values x_i from the data-generating process and assuming the sample mean is sufficient for μ (i.e., the sample mean covers all the relevant information in the data related to μ) the likelihood is proportional to Equation 7 (Wilks, 2011).

$$f(x|\mu) \propto \prod_{i=1}^{n} exp\left[-\frac{(x_i-\mu)^2}{2\sigma_h^2}\right] \tag{7}$$

Taking into account the sample mean is sufficient for μ (i.e., the sample mean covers all the relevant information in the data related to μ) the likelihood can be rephrased as Equation 8 (Wilks, 2011).

$$f(\bar{x}|\mu) \propto exp\left[-\frac{n(\bar{x}-\mu)^2}{2\sigma_*^2}\right] \tag{8}$$

Since *n* data are sampled from a Gaussian distribution with the parameters (μ, σ_*^2), they also have a Gaussian distribution with parameters $(\mu, \sigma_*^2/n)$, using Bayes' Theorem and combining prior and likelihood the posterior distribution will be as follow (Wilks, 2011):

$$f(\mu|\bar{x}) = \frac{1}{\sqrt{2\Pi}\sigma_h'} exp\left[-\frac{n(\mu-\mu_h')^2}{2\sigma_h^{2\prime}}\right] \tag{9}$$

where the posterior hyperparameters $(\mu_h', \sigma_h^{2\prime})$ are as follows (Wilks, 2011):

$$\mu_h' = \frac{{\mu_h}/{\sigma_h^2} + {n\bar{x}}/{\sigma_*^2}}{{1}/{\sigma_h^2} + {n}/{\sigma_*^2}} \tag{10}$$

$$\sigma_h^{2\prime} = \left(\frac{1}{\sigma_h^2} + \frac{n}{\sigma_*^2}\right)^{-1} \tag{11}$$

The posterior mean is indeed a weighted mean of prior and sample mean with much greater weight pertinent to the sample mean which rises as the sample size increases. This property leads to the less dependency of prediction on the less reliable old data and instead emphasizes recently sampled data. The variance of the posterior is also smaller than both the prior and the known data-generating variances and even decreases as the sample size increases. Another aspect regarding the posterior parameters is that since the variance of the data-generating process (σ_*^2) is known, only the sample mean appears in the estimations and neither the sample variance nor amount of additional data can enhance our knowledge about it.

The variability of the sampling, which is of a Gaussian data-generating process, combining with the uncertainty about μ which is expressed by posterior causes uncertainty about future values of x^+. Considering these two contributions, the Gaussian predictive distribution parameters (μ_+, σ_+^2) are as follows (Wilks, 2011):

$$\mu_+ = \mu_h' \tag{12}$$

$$\sigma_+^2 = \sigma_*^2 + \sigma_h^{2\prime} \tag{13}$$

2.3 Monte Carlo Simulation

MCS method is generally defined as a methodology to estimate the solution of mathematical problems by means of random numbers. Dealing with complex systems, MCS is known to be a powerful modeling tool to achieve close adherence to reality (Zio, 2013). MCS is a widely used modeling tool that its continuous improvement allows its application to complex systems and problems in a variety of scientific domains. One of the most important applications of MCS is in anticipating future patterns and determining how a system behaves based on the available information about the movements of the system among possible states. However, the prediction can never fit reality in all details (Zio, 2013).

For instance, in order to estimate an n-dimensional Euclidean volume V of a complex shape, V is to be placed inside a domain of a volume W that can be readily evaluated. By sampling a large number N of points inside W at random, n of these points will fall inside V, while the remaining $N - n$ will fall outside. Clearly, n is a random number that follows a binomial distribution with the parameter $p = \frac{V}{W}$ (i.e. the probability of which a sampled point falls inside the volume V). Consequently, considering n as an estimate of the average number of successes, the volume V is estimated as follows (Zio, 2013):

$$n \simeq Np = N\frac{V}{W} \text{ and } \hat{V} = \frac{n}{N}W \tag{14}$$

Similarly, the value of an integral can be estimated using a set of points randomly drawn from a distribution with support over the range of integration. Indeed, estimating the value of complicated integrals is the other common application of MCS. The value of a one-dimensional integral of the form $\int_a^b f(x)\,dx$ can be derived analytically for only a few functions f, whilst, numerical approximations of the integral are often considered as useful approaches for the rest functions. Particularly, dealing with Bayesian inference, an approximation of integrals is frequently required since prior or posterior distributions may not follow a familiar distributional family (Givens & Hoeting, 2013).

The problem of estimating the volume V is formally equivalent to the evaluation of a definite integral. In case the target density f is too complex to calculate its definite integral, MCS estimates its value using another density, g, so-called proposal density or envelope, from which is analytically easier to sample and covers f in its domain. Similar to the procedure mentioned to estimate an n-dimensional Euclidean volume V of a complex shape, here the N samples are randomly drawn from the proposal density. Thus, n of these points will fall inside f, while the remaining $N - n$ will fall outside. Consequently, the value of the integral over the range of interest can be estimated. For instance, instead of calculating the exact value of the integral for the Normal PDF, $N(\mu = 2, \sigma^2 = 2.25)$, in the range of $(1.5,2.5)$, alternatively, we drew $N = 1000$ random samples from the proposal function, $g(x) = 0.3$, in the same range. It should be mentioned here that the procedure of sampling from a density or distribution function in both continuous and discrete forms is later discussed in Section 2.4. Clearly, the share of each sample

is $\frac{0.3}{1000} = 0.0003$. Consequently, considering the samples that fell inside the normal function, which in our case is $n = 865$, the MCS estimation will be $865 \times 0.0003 = 0.260$. Then, the estimation has 0.001 deviations from the exact value, 0.261, that we already know. The process is depicted in Figure 5.

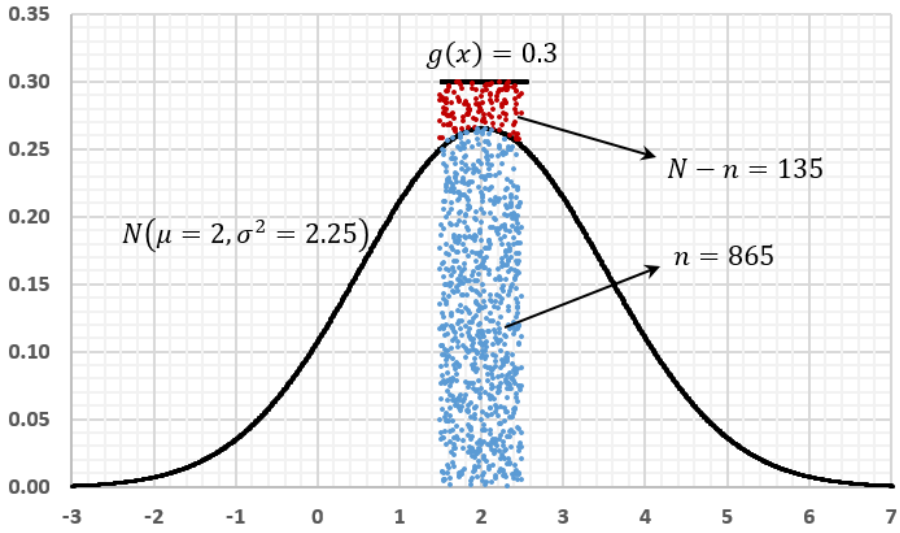


Figure 5. MCS method to estimate the value of the integral for Normal PDF, $N(\mu = 2, \sigma^2 = 2.25)$*, in the range of* $(1.5,2.5)$*, using* $N = 1000$ *random samples from the proposal function,* $g(x) = 0.3$

2.4 Inverse Transform Method of Sampling

Sampling is the foundation of the approximation methods in this study where a variety of discrete and continuous functions are to be used for random sampling in different algorithms. In this regard, the *inverse transform method* as a simple and commonly used sampling from density and distribution functions are used in this study (Givens & Hoeting, 2013).

2.4.1 Continuous distributions

Assuming the non-decreasing Cumulative Distribution Function (CDF), $F_X(x)$ is continuous and differentiable, the related PDF is:

$$f_X(x) = \frac{dF_X(x)}{dx}; \quad f_X(x) \geq 0; \quad \int_{-\infty}^{+\infty} f_X(x)dx = 1 \tag{15}$$

Therefore, the value of the integral of $f_X(x)$ over the interval, Δx can be estimated by sampling a sequence of $N \gg 1$ values from $F_X(x)$. Then, we have:

$$\frac{n}{N} \simeq \int_{\Delta x} f_X(x)dx \tag{16}$$

where n is the number of sampled points falling within the interval Δx (see Section 2.3).

According to the inverse transform method, in order to simulate random draws from a target density $f_X(x)$, the corresponding CDF is used. Given that, X is a random variable obeying the CDF $F_X(x)$, denoted as $X \sim F_X(x)$, we have:

$$P(X \leq x) = \int_{-\infty}^{x} f_X(x)dx = F_X(x); \quad F_X(-\infty) = 0; \quad F_X(+\infty) = 1 \tag{17}$$

Therefore, taking into account that CDF is a non-decreasing function over the interval [0,1), randomly sampled value, r, from the uniform distribution in the interval [0,1), denoted as $R \sim U(0,1)$, the corresponding value from the density $f_X(x)$ obtains from the inverse of $F_X(x)$ as follows:

$$F_X^{-1}(r) = \inf\{x: F_X(x) \geq r\} \tag{18}$$

where $\inf\{.\}$ is the infimum function, which indicates the value x relating to the smallest value of $F_X(x)$ that is greater than the lower bound r. In other words,

$$x = F_X^{-1}(r) \tag{19}$$

and

$$P(R \leq r) = P(X \leq x) \tag{20}$$

$$U_R(r) = F_X(x) \tag{21}$$

$$r = \int_{-\infty}^{x} f_X(x)\, dx \tag{22}$$

where $U_R(r)$ indicates that r is from $R \sim U(0,1)$. The procedure is depicted in Figure 6. Accordingly, for any target density with invertible CDF, finding the value x corresponding the

value r, randomly sampled from $U(0,1)$, is equivalent to randomly sampling from the target density $f_X(x)$.

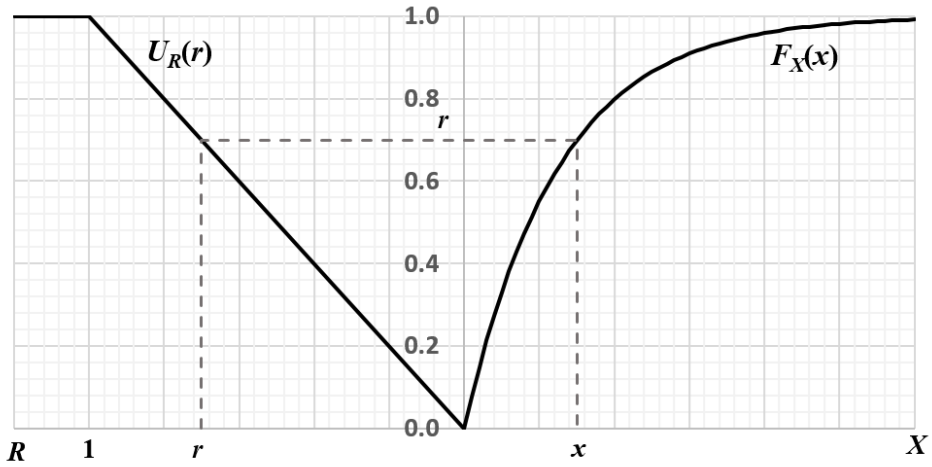


Figure 6. Inverse transform method to sample from a continuous distribution, adapted from (Zio, 2013)

2.4.2 Discrete distributions

Considering X as a random variable, which can only hire discrete values x_k, with probabilities

$$f_k = P(X = x_k) \geq 0;\ k = 0,1,2, \ldots \tag{23}$$

and ordering the sequence of $\{x\}$ so that $x_{k-1} < x_k$, the CDF is then

$$F_k = P(X \leq x_k) = \sum_{i=0}^{k} f_i = F_{k-1} - f_k; k = 0,1,2, \ldots \tag{24}$$

where it is assumed that $F_{-1} = 0$. Clearly, we have

$$\lim_{k\to\infty} F_k = 1 \tag{25}$$

Given that r is randomly sampled from $U(0,1)$, the probability that r falls in the interval $(F_{k-1}, F_k]$ is

$$P(F_{k-1} < r \leq F_k) = \int_{F_{k-1}}^{F_k} dr = F_k - F_{k-1} = f_k = P(X = x_k) \tag{26}$$

The procedure of sampling from discrete CDF using the continuous PDF, $U_R(r)$, is indicated in Figure 7.

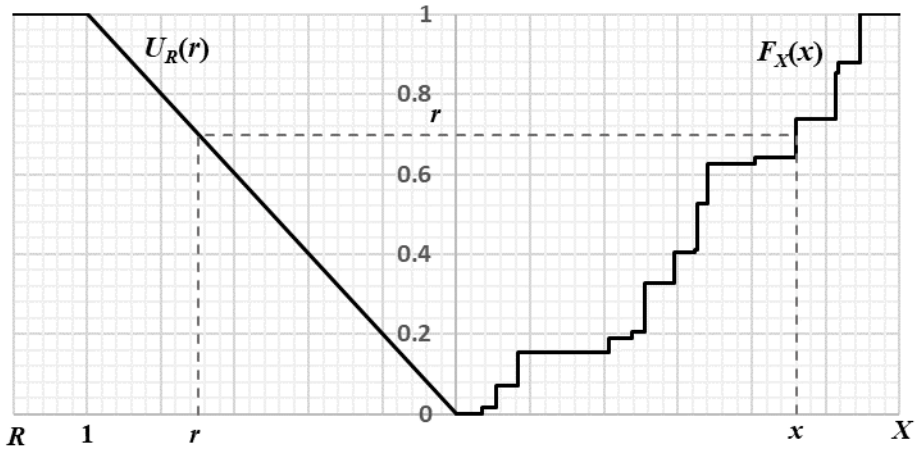


Figure 7. Inverse transform method to sample from a discrete distribution

An example of inverse transform method to sample from a discrete distribution is illustrated in Figure 8 in which it is assumed that r falls within interval $(F_1, F_2]$. Moreover, the steps of the algorithm are shown in Figure 9.

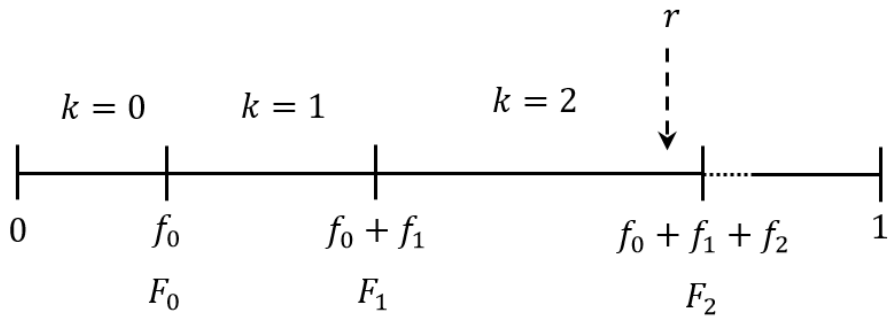


Figure 8. Inverse transform method example to sample from a discrete distribution, adapted from (Zio, 2013)

1. *Sample randomly from* $U(0,1)$
2. *Set* $k = 0; F = f_0$
3. *If* $r \leq F$ *proceed to 5*
4. *If* $r > F$, *set* $k \leftarrow k + 1; F \leftarrow F + f_k$ *and proceed to 3*
5. $X = x_k$

Figure 9. The inverse transform method algorithm for sampling from discrete distributions (Zio, 2013)

2.5 Rejection Sampling

As mentioned, in order to estimate the value of integral for complex density functions, MCS samples randomly from a proposal density g that is analytically easier to sample. However, to make the procedure faster and more efficient and avoid wasting the memory with useless draws, *rejection sampling* obtains a random draw from exactly the target distribution. Accordingly, while sampling from g, the sampling probability is being corrected through random rejection of some candidates (Givens & Hoeting, 2013).

Considering the proposal density g that is easier to sample and calculate $g(x)$, and $e(x)$ as an envelope with the property of $e(x) = g(x)/\alpha \geq f(x)$ for all x for which $f(x) > 0$ for a given constant $\alpha \leq 1$. The rejection sampling algorithm is depicted in Figure 10 (Givens & Hoeting, 2013).

1. *Sample* $y \sim g$.
2. *Sample* $r \sim U(0,1)$.
3. *Reject* y *if* $r > f(y)/e(y)$ *and then return to step 1. Otherwise, proceed 4.*
4. *Keep the value of* y. *Set* $x = y$, *and consider* x *to be an element of the target random sample.*
5. *Stop if the sample of the desired size has been accumulated. Otherwise, return to step 1.*

Figure 10. Rejection sampling algorithm (Givens & Hoeting, 2013)

As proved by Givens and Hoeting (2013), the kept draws using this algorithm are i.i.d. samples from the target density f and no approximation is involved. Meanwhile, α can be interpreted as the expected proportion of accepted candidates. Thus, α is a measure of the efficiency of the algorithm by influencing the number of iterations. As illustrated in Figure 11, the samples were drawn from the shaded region under envelope $e(x)$ and above $f(x)$ are to be rejected and therefore, envelopes that only marginally exceed f produce fewer wasted (i.e. rejected) draws which correspond to α values close to 1 (Givens & Hoeting, 2013).

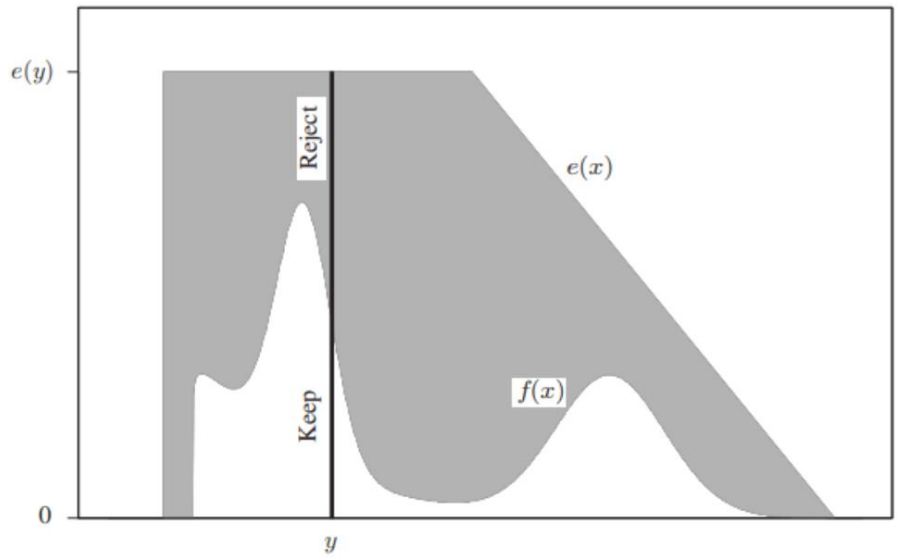


Figure 11. Rejection sampling for a target density f(x) using envelope e(x) (Givens & Hoeting, 2013)

2.6 Importance Sampling

In data analysis, summarizing the results in the forms of expectation such as marginal mean, variance, and covariance is often of interest. The expected value of the quantity of interest, $h(\theta)$, is computed as below (Ridgeway & Madigan, 2003):

$$E(h(\theta)|x_1, \dots, x_N) = \int h(\theta) f(\theta|x_1, \dots, x_N) d\theta \tag{27}$$

where $f(\theta|X)$, is the posterior density of the parameters given the observed data. However, except the simple examples, the computation of these integrals is difficult in closed form. As previously mentioned, Monte Carlo integration methods estimate these integrals by sampling from the posterior, $f(\theta|X)$, and appealing to the law of large numbers. Accordingly, the expected value of the quantity of interest, $h(\theta)$, is approximated as below (Ridgeway & Madigan, 2003):

$$\lim_{M\to\infty} \frac{1}{M} \sum_{i=1}^{M} h(\theta_i) = \int h(\theta) f(\theta|x_1, \dots, x_N) d\theta \tag{28}$$

where the θ_i compose a sample from $f(\theta|X)$.

Importance Sampling (IS) is a relatively more efficient form of the MCS method to approximate integrals, particularly when sampling from the "target density," $f(\theta|X)$ is not readily available. In this case, IS draws from another available "sampling density," $g(\theta)$ that is easy to sample, also known as "proposal density" or "envelope," and approximates the integral as follows (Ridgeway & Madigan, 2003; Givens & Hoeting, 2013):

$$\int h(\theta)f(\theta|x_1,\dots,x_N)d\theta = \int h(\theta)\frac{f(\theta|X)}{g(\theta)}g(\theta)d\theta = \lim_{M\to\infty}\frac{1}{M}\sum_{i=1}^{M} w_i h(\theta_i) \tag{29}$$

where θ_i is drawn from $g(\theta)$ and $w_i = f(\theta_i|X)/g(\theta_i)$. Considering that the expected value of w_i under $g(\theta)$ is 1, the only thing needed is to compute weights up to a constant of proportionality and then normalize (Ridgeway & Madigan, 2003):

$$\int h(\theta)f(\theta|x_1,\dots,x_N)d\theta = \lim_{M\to\infty}\frac{\sum_{i=1}^{M} w_i h(\theta_i)}{\sum_{i=1}^{M} w_i} \tag{30}$$

2.7 Sampling Importance Resampling

Briefly, in the IS method, samples are drawn from a proposal density, g, and are weighted to correct the sampling probabilities so that the weights are related to a target density f. The weighted sample is particularly useful to estimate expectations under f. The weights can also be standardized so they sum to 1, although it is not necessary. Therefore, IS can be seen as an approximation of f by a discrete distribution and weights as masses of observed points. Rubin (1987, 1988) proposed sampling from this discrete distribution, which is called *Sampling Importance Resampling* (SIR). Accordingly, as the number of samples increases, the distribution of the random draws converges to f (Givens & Hoeting, 2013).

Comparing SIR with rejection sampling, both SIR and rejection sampling rely on the ratio of target to the envelope, while they differ in the number of draws. Whereas a pre-determined number of draws is required in SIR to generate a sample of size *n*, in rejection sampling the number of draws for the same sample size is random. Moreover, the distribution of a generated draw by rejection sampling is exactly f, while the SIR algorithm permits a random degree of approximation to f in the distribution of the sampled points (Givens & Hoeting, 2013).

2.8 Sequential Monte Carlo

The efficiency of SIR declines and it can be difficult to implement when the target density f becomes high dimensional. It is challenging to specifying a very good high-dimensional envelope, g, that properly approximates f with sufficiently heavy tails but little waste. This drawback is addressed by *Sequential Monte Carlo* (SMC) methods according to which the high-dimensional task is splitting into a sequence of simpler steps, each of which updates the previous one (Givens & Hoeting, 2013).

Let $X_{1:t} = (X_1, \dots, X_t)$ denotes a discrete-time stochastic process where X_t is the observation at time t and $X_{1:t}$ represents the entire history of the sequence thus far. For simplicity, the scalar notation is adopted here; however, X_t may be multidimensional. Meanwhile, the density of $X_{1:t}$ is denoted as f_t. Consider that at time t the expected value of $h(X_{1:t})$ is supposed to be estimated with respect to f_t and using an IS strategy (Givens & Hoeting, 2013).

One strategy would be directly used of the SIR approach to sample $X_{1:t}$ sequences from an envelope g_t and then the expected value of $h(X_{1:t})$ can be estimated by calculating the importance weighted average of this sample. However, in this strategy as t is increasing, $X_{1:t}$ and the expected value of $h(X_{1:t})$ evolve. Therefore, at the time t it would be reasonable to update previous inferences rather than acting in a way that there is no previous information. An alternative strategy is to append the simulated X_t to the $X_{1:t-1}$ that previously simulated. Consequently, to estimate the expected value of $h(X_{1:t})$, the previous importance weights are to be adjusted. This approach is called *Sequential Importance Sampling* (SIS) (Liu & Chen, 1988).

2.8.1 Sequential Importance Sampling for Markov Processes

Assuming $\mathbf{X}_{1:t}$ is a Markov process, X_t depends only on $\mathbf{X}_{t-1}$ rather than the whole history $\mathbf{X}_{1:t-1}$. Accordingly, the target density $f_t(\mathbf{x}_{1:t})$ can be expressed as follows (Givens & Hoeting, 2013):

$$\begin{aligned} f_t(\mathbf{x}_{1:t}) &= f_1(x_1) f_2(x_2|\mathbf{x}_{1:1}) f_3(x_3|\mathbf{x}_{1:2}) \dots f_t(x_t|\mathbf{x}_{1:t-1}) \\ &= f_1(x_1) f_2(x_2|x_1) f_3(x_3|x_2) \dots f_t(x_t|x_{t-1}) \end{aligned} \quad (31)$$

Similarly, by adopting the same Markov form for the envelope, we have (Givens & Hoeting, 2013)

$$g_t(\mathbf{x}_{1:t}) = g_1(x_1)g_2(x_2|x_1)g_3(x_3|x_2) \dots g_t(x_t|x_{t-1}) \tag{32}$$

According to the ordinary non-sequential SIR, as a sample is drawn from $g_t(\mathbf{x}_{1:t})$ at time t, each $\mathbf{x}_{1:t}$ is to be reweighted by $w_t = f_t(\mathbf{x}_{1:t})/g_t(\mathbf{x}_{1:t})$. Whilst, based on SIS in a Markov process we have (Givens & Hoeting, 2013):

$$w_t = u_1 u_2 \dots u_t \tag{33}$$

where $u_1 = f_1(x_1)/g_1(x_1)$ and $u_i = f_i(x_i|x_{i-1})/g_i(x_i|x_{i-1})$ for $i = 2,3, \dots, t$.

Having $\mathbf{x}_{1:t-1}$ and w_{t-1} in hand and using the Markov property the next component, X_t, can be sampled and appended to $\mathbf{x}_{1:t-1}$. Moreover, w_{t-1} can be adjusted using the multiplicative factor u_t. The SIS for Markov processes at time t is given in the steps below. Using a sample of n such points and their weights, $f_t(\mathbf{x}_{1:t})$ and thus the expected value of $h(\mathbf{X}_{1:t})$ can be approximated. The algorithm is given in Figure 12 (Givens & Hoeting, 2013).

1. *Sample* $X_1 \sim g_1$. *Let* $w_1 = u_1 = f_1(x_1)/g_1(x_1)$. *Set* $t = 2$.
2. *Sample* $X_t|x_{t-1} \sim g_t(x_t|x_{t-1})$.
3. *Append* x_t *to* $\boldsymbol{x}_{1:t-1}$, *obtaining* $\boldsymbol{x}_{1:t}$.
4. *Let* $u_t = f_t(x_t|x_{t-1})/g_t(x_t|x_{t-1})$.
5. *Let* $w_t = w_{t-1}u_t$. *At the current time,* w_t *is the importance weight for* $\boldsymbol{x}_{1:t}$.
6. *Increment* t *and return to step 2.*

Figure 12. Sequential importance sampling algorithm (Givens & Hoeting, 2013)

To obtain an independent sample of size n from $\mathbf{X}_{1:t}^{(i)}, i = 1,2, \dots, n$, the algorithm in Figure 12 can be carried out considering the n sequences one at a time or as a batch. Consequently, the estimation for the weighted average of $h(\mathbf{X}_{1:t})$ is as below (Givens & Hoeting, 2013):

$$E_{f_t}[h(\mathbf{X}_{1:t})] = \sum_{i=1}^{n} w_t^{(i)} h\left(\mathbf{X}_{1:t}^{(i)}\right) / \sum_{i=1}^{n} w_t^{(i)} \quad (34)$$

The standardization of the weights at the end of each cycle is not essential, while if the estimation of $E_{f_t}[h(\mathbf{X}_{1:t})]$ is of interest, the normalization is natural (Givens & Hoeting, 2013).

2.9 Markov Chain Monte Carlo

MCMC is considered to be the most common computational method for Bayesian analysis of complex models. Whereas IS generates independent draw and related weights, MCMC methods build a Markov chain, generating dependent draws that have stationary density. Although creating such a Markov chain is often easy, there is still a bit of art required to construct an efficient chain with reasonable convergence speed. (Ridgeway & Madigan, 2003; Givens & Hoeting, 2013)

2.9.1 The Metropolis-Hastings Algorithm

A very general method to implement MCMC is the Metropolis-Hastings algorithm. Given that we already have a sample θ_1 from a target density $f(\theta|x)$, a new sample θ' is drawn from a proposal density $g(\theta|\theta_1)$. Moreover, the new sample θ' is accepted or rejected according to an acceptance probability, which depends on the previous draw and is to be updated in each iteration. Then, one of the key properties of this method is that the density of θ_2 will also be $f(\theta|x)$. The algorithm obtains a sequence $\theta_1, \theta_2, \dots, \theta_M$ with the stationary density of $f(\theta|x)$ (Ridgeway & Madigan, 2003; Givens & Hoeting, 2013). The algorithm is shown in Figure 13.

1. *Initialize* θ_1
2. *For* $i = 2,3,\dots,M$ *do*
   - 2.1. *Sample a* $\theta' \sim g(\theta|\theta_{i-1})$
   - 2.2. *Compute the acceptance probability* $\alpha(\theta', \theta_{i-1}) = min\left\{1, \frac{f(\theta'|x)g(\theta_{i-1}|\theta')}{f(\theta_{i-1}|x)g(\theta'|\theta_{i-1})}\right\}$
   - 2.3. *With probability of* $\alpha(\theta', \theta_{i-1})$ *set* $\theta_i = \theta'$. *Otherwise, set* $\theta_i = \theta_{i-1}$

Figure 13. The Metropolis-Hastings algorithm (Ridgeway & Madigan, 2003)

2.9.2 Convergence Diagnostic

One of the most critical issues associated with the implementation of MCMC simulation is the time before the chain settles down to a steady-state so-called "converged". To mitigate the possibility of bias due to the effect of the starting values, the iterations within the initial transient phase are usually discarded. Rates of convergence on different target distributions vary considerably, which makes it difficult to determine the length of the required initial transient phase (Brooks & Roberts, 1998). Many techniques have been developed trying to determine the convergence of a particular Markov chain as reviewed by Brooks and Roberts (1998). However, it is not generally possible to estimate the Markov chain convergence rate and then determine sufficient iterations to satisfy a prescribed accuracy measure (Tierney, 1994; Cowles & Carlin, 1996).

In this study, the variability of the estimations in the iterations is evaluated via the absolute form of Coefficient of Variation (CV) and considered as a useful measure of convergence. To this aim, the algorithm is firstly supposed to proceed with a certain amount of iterations and then continues as long as the last m outcomes considerably vary (Wilks, 2011; Zio, 2013). Thus, the absolute value of CV for the last m outcomes, CV_m, which is calculated below, should be larger than a threshold, let say CV^T.

$$|CV_m(X)| = \left|\frac{S(X)}{E(X)}\right| \tag{35}$$

where $S(X)$ and $E(X)$ indicate standard deviation expected value of the sampled values, respectively.

2.10 MINCOG Icing model

Sea spray is generated during the collision of vessel and waves, and by strong winds ripping off small droplets from the crest of breaking waves. However, the amount of water generated by wind is much smaller compared to sea spray generated during the collision of vessel and waves. The collision of waves and vessels leads to the formation of spray-cloud, which its droplets will be transported by air and settled onto different surfaces of the vessel (see Figure 14). In this regard, waves, vessel, seawater, and air contribute to the ice generation.

Accordingly, the characteristic of each factor such as height and period of wave, speed and direction of the vessel, the salinity of seawater, wind speed, temperature, relative humidity, and pressure of air are influential (Kulyakhtin and Tsarau, 2014; Samuelsen, Edvardsen, and Graversen, 2017). In this study, the prediction of some of the parameters including wave height, wind speed, temperature, relative humidity, atmospheric pressure, and wave period, is purposed. Then, the predicted parameters are considered as input for the MINCOG model. Meanwhile, the Norwegian Coast Guard ship class, KV Nordkapp, is considered by the MINCOG model as a reference ship type for the vessel-icing calculations (see Figure 15). More details about the MINCOG model and the vessel can be found in (Samuelsen, Edvardsen, & Graversen, 2017).

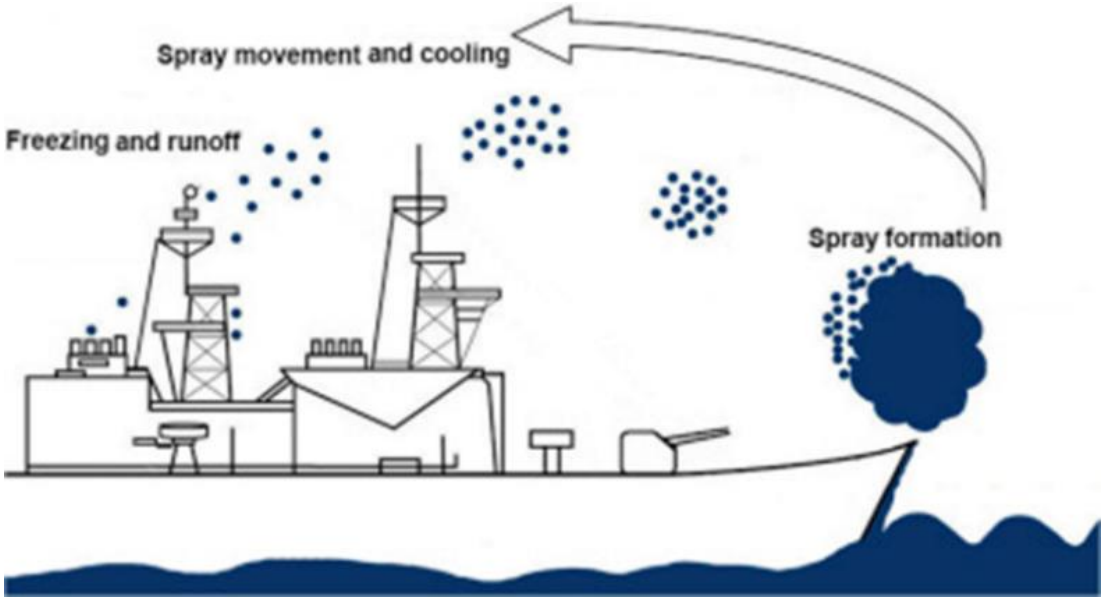


Figure. 14. Schematic of the spray formation and movement on a ship (Dehghani-sanij, Muzychka, & Naterer, 2015)

Figure. 15. KV Nordkapp class vessel (Samuelsen, Edvardsen, and Graversen, 2017)

3 Experiments

As mentioned before, the scope of the study is the Arctic offshore, specifically, the sea area between Northern Norway and Svalbard archipelago bounded to the latitudes 69°N to 78°N and longitudes 8°E to 36°E (see Figure 16). The offshore location with coordinate (74.07°N, 35.81°E) is selected for analysis of the meteorological and oceanographic parameters and, icing events and rates since is currently open for petroleum activity. This coordinate, as illustrated in Figure 16, is located approximately 500 km east of Bjørnøya in the Norwegian part of the Barents Sea, where the discovery wellbore 7435/12-1 was drilled in 2017 (Norwegian Petroleum Directorate, 2020).

Reanalysis data from NORA10 during 33 years are applied to evaluate the performance of the models. The dataset involves 3-hourly data from 1 January 1980 to 31 December 2012 provided by the Norwegian Meteorological Institute (MET). The data consist of meteorological and oceanographic parameters including wave height, wind speed, temperature, relative humidity, atmospheric pressure and wave period, and icing events (Reistad, et al., 2011).

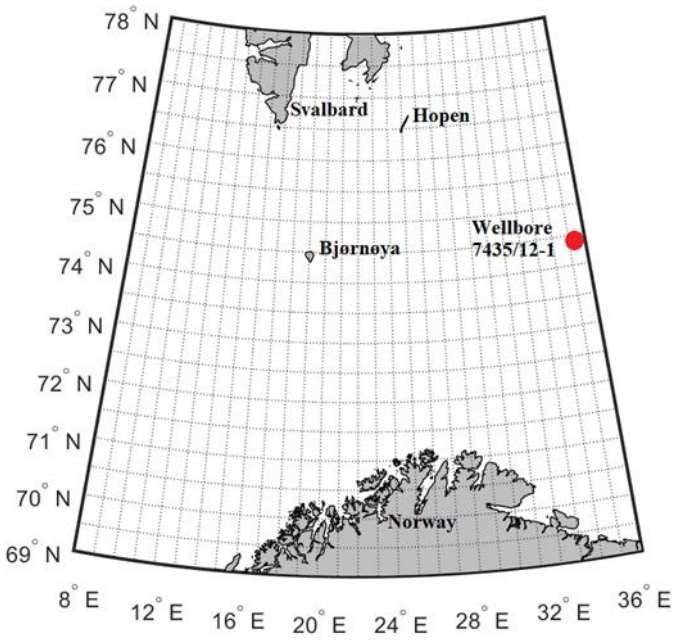


Figure 16. The study area between Northern Norway and Svalbard archipelago bounded to the latitudes 69°N to 78°N and longitudes 8°E to 36°E, adapted from (Naseri & Samuelsen, 2019).

Obviously, the models and algorithms mentioned in the methodology cannot be directly used in their standard form and some modifications and assumptions are required. For instance, in the Bayesian approach, to consider the effect of climate change, the old less reliable data are used to estimate the prior distribution, which is modified by the newer sets of data to estimate the posterior distribution. Moreover, the Normal distribution function is used in the Bayesian approach to fit the data. Whilst, in SIS and MCMC models the data are fitted using kernel smoothing function, which is indeed the target density, and the Weibull distribution function is used as the proposal density.

In the following, after defining required assumptions, proposed models are developed by combining and modifying the algorithms that were mentioned in the methodology. Thereafter, using the data over 32 years (1980-2011), the six meteorological and oceanographic parameters are simulated for one year on a daily basis and the results are compared with the 33rd year (2012). Then, the predicted data are used as input parameters in the MINCOG model to forecast the icing rate.

3.1 Proposed Bayesian approach

In the proposed Bayesian framework, to mitigate the effect of climate change, the old less reliable data, during 27 years from 1980 to 2006, are used to estimate the prior distribution, which is modified by the newer sets of data, from 2007 to 2011, to estimate the posterior distribution. Accordingly, it is assumed that the prior data-generating process for the daily average of each parameter is Gaussian over the years 1980 to 2006. This assumption is evaluated for the six parameters in coordinate (74.07°N, 35.81°E) using the Anderson-Darling test at the 5% significance level, where the null hypothesis (i.e., H0) is that the parameter is from a population with a Normal distribution, against the alternative hypothesis (i.e., H1) that the parameter is not from a population with a Normal distribution (Anderson & Darling, 1952). Accordingly, and based on the data over 27 years (1980-2006), the null hypothesis cannot be rejected in the majority of the days. The results of the Anderson-Darling test are shown in Table 3, where the number and percentage of the days that the null hypothesis cannot be rejected are mentioned. Likewise, the posterior data-generating process is assumed to be Gaussian with unknown parameters, and then the newer sets of data during five years (2007-2011) are

considered as a sample to modify the prior distribution and determine the posterior distribution. However, the variance of the data-generating process, σ_*^2, is considered to be known and daily average values of parameters during 32 years (1980-2011) are used to calculate this variance.

Table 3. The Anderson-Darling test at the significance level of 5%, for normality test of the daily average of parameters in coordinate (74.07°N, 35.81°E)

Parameter	Number of days in year which H0 cannot be rejected	Percentage of days in year which H0 cannot be rejected
Wave height	245	67 %
Wind speed	330	90 %
Temperature	257	70 %
Relative humidity	284	78 %
Atmospheric pressure	346	95 %
Wave period	180	49 %

3.2 Proposed SIS

The drawback of the standard SIS that defined in Section 2.8.1 is that the deviation of estimation at each state will be added to the deviation from the previous state so that the algorithm hardly converges to a value. Therefore, rather than sampling dependent draws from static densities, the target and proposal densities are defined in a way to be dependent on the previous state. To this aim, the deviation of each day from the previous day in 32 years (1980-2011) was extracted from the dataset. Therefore, the conditional density at each state is achieved by estimating the average of the previous state adding to the possible deviations. Accordingly, a kernel smoothing density with Gaussian kernel function (see Table 1) is considered as the target density. Moreover, Weibull distribution is used as the proposal density. However, since Weibull distribution is defined only for positive values, a data shifting procedure is embedded in the algorithm. Accordingly, a positive value, A, is to be added to all the data, which is calculated via equation 36.

$$A = \left|\min_z\{x_z\} + 1\right| \tag{36}$$

where z is the index of bins in the kernel density estimate and x_z represents the center of the z^{th} bin. It should be mentioned that 'A' must be later subtracted from the simulated results.

Furthermore, the parameters of the Weibull distribution are estimated by the Maximum Likelihood Estimation (MLE) method considering 32 years of the data (1980-2011). Consequently, the algorithm iterates until a defined number of samples, M, is drawn. Additionally, the performance of the algorithm using two sizes of $M = 200$ and $M = 500$, namely SIS200 and SIS500, are investigated. The algorithm is outlined in Figure 17.

For *each parameter,* θ, **do**
  For *each day* $d = 1, .., D$ **do**
    For *each year* $y = 1, .., Y$ **do**
      Load *the related historical data,* x, *from dataset*
      Estimate *the target density,* $f_d(\theta_d) \sim kernel\ density\ using\ x$
      Calculate $A = \left| \min_k \{x_k\} + 1 \right|$; *where* x_k *is the center of bin* k^{th} *in the kernel estimation*
      For $i = 1, .., N$ **do**
        $x_i \leftarrow x_i + A$
      End
      Let *the proposal density,* $g_d(\theta_d) \sim Weibull(a, b)$, *and estimate the parameters using MLE method and new* x
      Let *the daily mean value of new* x *as the initial sample,* $\theta_{d,1}$
      Calculate *the related importance sampling weight,* w_1
      Set *iteration lower bound,* M
      Set *CV threshold,* CV^T
      For $j = 2, \ldots, M$ **do**
        Sample $\theta_{d,j}$ *using Metropolis-Hastings algorithm and calculate* w_j
        Calculate *the weighted average of all drawn samples,* $IS_j(\theta_d)$
      End
      While $|CV_m(IS)| > CV^T$ **do**
        $j \leftarrow j + 1$
        Sample θ_j *using Metropolis-Hastings algorithm and calculate* w_j
        Calculate *the weighted average of all drawn samples,* $IS_j(\theta_d)$
      End
      Let $IS_j(\theta_d) - A$ *value as the estimation of the parameter*
    End
  End
End

Figure 17. Proposed SIS algorithm

3.3 Proposed MCMC

To apply the MCMC approach, a modified Metropolis-Hastings algorithm is developed in which the concepts of rejection sampling and IS are embedded. Similar to SIS, the kernel smoothing density and Weibull distribution function is considered as the target and proposal

density, respectively. Furthermore, a dynamic stopping criterion is added to the algorithm based of which the algorithm iterates until the CV (see equation 35) in the last $m = 50$ iterations drops below CV threshold, $CV^T = 0.01$, that implies the algorithm no longer achieves different results. Moreover, to avoid early stoppage, the stopping criterion is to be activated after a certain amount of iterations, M, namely *iteration lower bound*. Thus, two sizes of $M = 200$ and $M = 500$ are later investigated which are called MCMC200 and MCMC500, respectively. Figure 18 indicates the steps of the algorithm.

For *each parameter,* θ, **do**
  For *each day* $d = 1, .., D$ **do**
    For *each year* $y = 1, .., Y$ **do**
      Load *the related historical data,* x, *from dataset*
      Estimate *the target density,* $f_d(\theta_d) \sim kernel\ density$ *using* x
      Calculate $A = \left| \min_k \{x_k\} + 1 \right|$; *where* x_k *is the center of bin* k^{th} *in the kernel estimation*
      For $i = 1, .., N$ **do**
        $x_i \leftarrow x_i + A$
      End
      Let *the proposal density,* $g_d(\theta_d) \sim Weibull(a, b)$, *and estimate the parameters using MLE method and new* x
      Let *the daily mean value of new* x *as the initial sample,* $\theta_{d,1}$
      Calculate *the related importance sampling weight,* w_1
      Set *iteration lower bound,* M
      Set *CV threshold,* CV^T
      For $j = 2, \ldots, M$ **do**
        Sample $\theta_{d,j}$ *using Metropolis-Hastings algorithm and calculate* w_j
        Calculate *the weighted average of all drawn samples,* $IS_j(\theta_d)$
      End
      While $|CV_m(IS)| > CV^T$ **do**
        $j \leftarrow j + 1$
        Sample θ_j *using Metropolis-Hastings algorithm and calculate* w_j
        Calculate *the weighted average of all drawn samples,* $IS_j(\theta_d)$
      End
      Let $IS_j(\theta_d) - A$ *value as the estimation of the parameter*
    End
  End
End

Figure 18. Proposed MCMC algorithm

3.4 Results and Discussion

Considering the aforementioned assumptions, the proposed models are programmed in MATLAB R2020a and run on a 1.60 GHz Intel® Core™ i5-8265U CPU and 8GB of RAM. Then, the meteorological and oceanographic parameters (i.e., wave height, wind speed, temperature, relative humidity, atmospheric pressure, and wave period) are predicted using a 32-year set of data from 1980 to 2011 in coordinate (74.07°N, 35.81°E) for all the days of 2012.

Moreover, the elements of the Bayesian inference comprised of the prior, posterior, sample, and predictive distributions are evaluated. As an example, the results related to the daily average temperature of 1st of April are shown in Table 4 in which $(\mu_{re}, \sigma_{re}^2)$ indicate mean and variance of reanalysis values in 2012. Moreover, the prior, sample, posterior, predictive, and the related reanalysis distributions are depicted in Figure 19. Apparently, the sample distribution is relatively closer to the reanalysis distribution rather than the prior distribution. Additionally, the Gaussian predictive distribution is more analogous to the reanalysis distribution in terms of central tendency as well as deviation. Therefore, making decisions for the future based on such a predictive distribution seems to be much more reliable than counting on the prior belief.

Table 4. Bayesian inference elements for temperature in coordinate (74.07°N, 35.81°E) on 1st of April

Parameter	Value
σ_*^2	1.12
(μ_h, σ_h^2)	(-3.49, 10.29)
$\bar{x}$	(-5.20, 8.52)
$(\mu_h', \sigma_h^{2\prime})$	(-5.16, 0.25)
(μ_+, σ_+^2)	(-5.16, 1.50)
$(\mu_{re}, \sigma_{re}^2)$	(-4.44, 0.95)

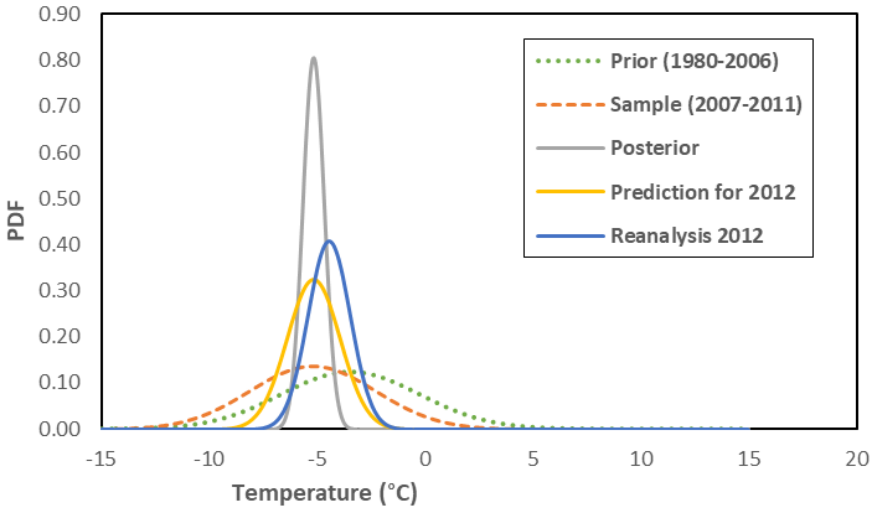


Figure 19. Bayesian inference elements for daily average temperature in coordinate (74.07°N, 35.81°E) on 1 April 2012

Additionally, to demonstrate how the introduced MCMC algorithm can cope with the estimation of complex empirical densities, two examples related to daily averages of relative humidity and wave height in coordinate (69.3°N, 8.6°E) are illustrated in Figures 20 and 21, respectively. The reason for choosing relative humidity and wave height and coordinate (69.3°N, 8.6°E), which is different from the location of the discovery wellbore 7435/12-1, is that these parameters represent relatively more complex behaviors in this coordinate. Thus, they can better challenge the performance of the model.

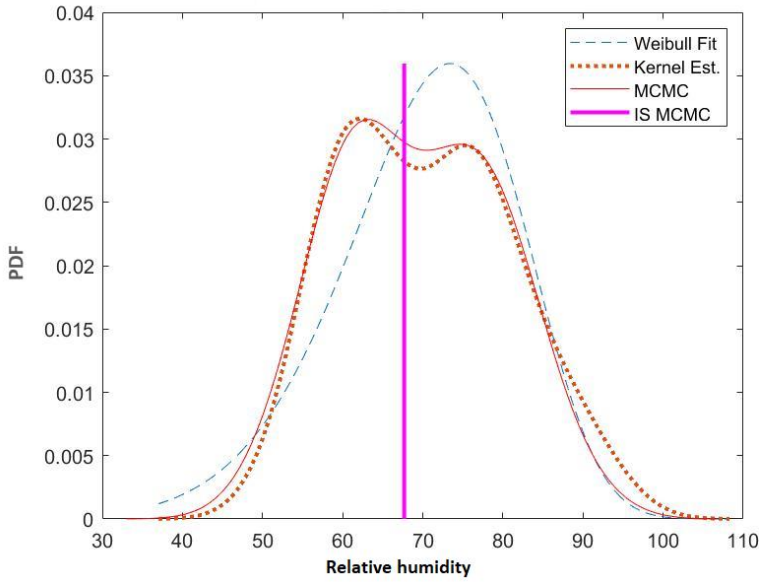


*Figure 20. Comparing target (kernel estimation), proposal (Weibull distribution), MCMC samples and estimation of daily average (IS MCMC) for relative humidity (100*fraction) in coordinate (69.3°N, 8.6°E) on 19 January 2012*

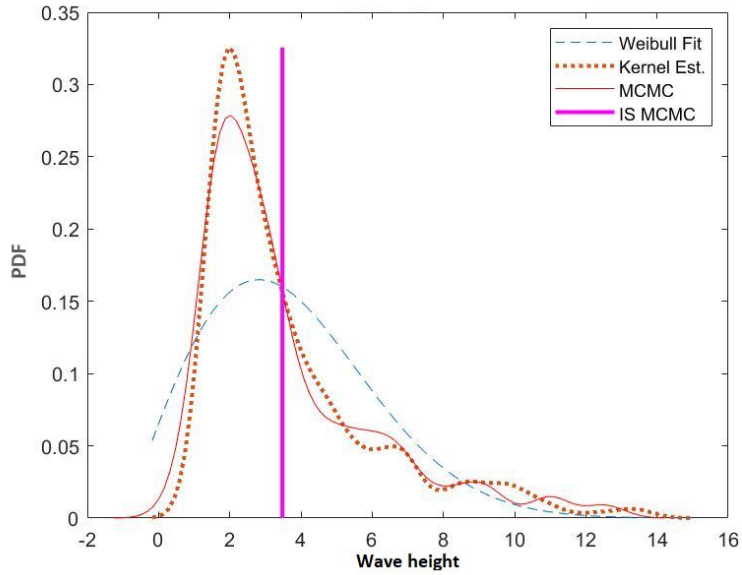


Figure 21. Comparing target (kernel estimation), proposal (Weibull distribution), MCMC samples and estimation of daily average (IS MCMC) for wave height (m) in coordinate (69.3°N, 8.6°E) on 8 January 2012

Furthermore, to show the efficiency of the defined stopping criteria in the proposed MCMC, the convergence of the model with the iteration lower bound of $M = 200$, as indicated by MCMC200, to simulate wind speed in coordinate (74.2°N, 14.6°E) on 7 February 2012 is depicted in Figure 22. Accordingly, the simulation procedure stops at the iteration 205 and avoids wasting time on further iterations that might lead to very small changes.

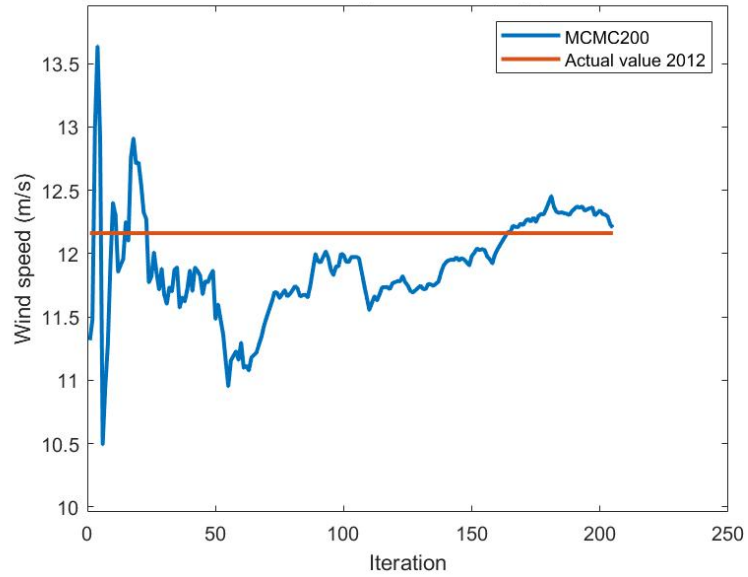


Figure 22. Convergence ratio of the MCMC with iteration lower bound of 200 to simulate wind speed in coordinate (74.2°N, 14.6°E) on 7 February 2012

The monthly averages of the predicted parameters from different algorithms are compared with the reanalysis values in 2012, as shown in Figure 23-28. Accordingly, there is no significant difference between the techniques and all of them demonstrate proper performance dealing with simulating meteorological and oceanographic parameters in the study area in the Arctic. However, the estimates of the Bayesian approach is slightly closer to the reanalysis values. Furthermore, examining different number of iterations (i.e. $M = 200$ and $M = 500$) revealed the capabilities of both SIS and MCMC algorithms to simulate the parameters with a relatively low number of iterations while further iterations result in only little improvements in some cases. Besides, the monthly averages of deviations from reanalysis values for different parameters are illustrated in Tables 5-10. Thus, the closest estimates to reanalysis values for wave height, wind speed, temperature, relative humidity, atmospheric pressure, and wave period are achieved by SIS200, MCMC500, MCMC500, MCMC500, SIS500, and Bayesian approach with deviations of 0.38 m, 1.53 m/s, 0.57 °C, 4.88%, 4.2 hPa, and 0.6 s, respectively.

Considering the combination of 5 simulation techniques and 12 months, we have 72 scenarios of which in 35 scenarios Bayesian approach has resulted in the lowest deviation from reanalysis value. This amount is 7, 10, 9, and 11 for SIS200, SIS500, MCMC200, and MCMC500, respectively. Therefore, the Bayesian approach is the most resistant technique, which is also robust due to few required assumptions for implementation.

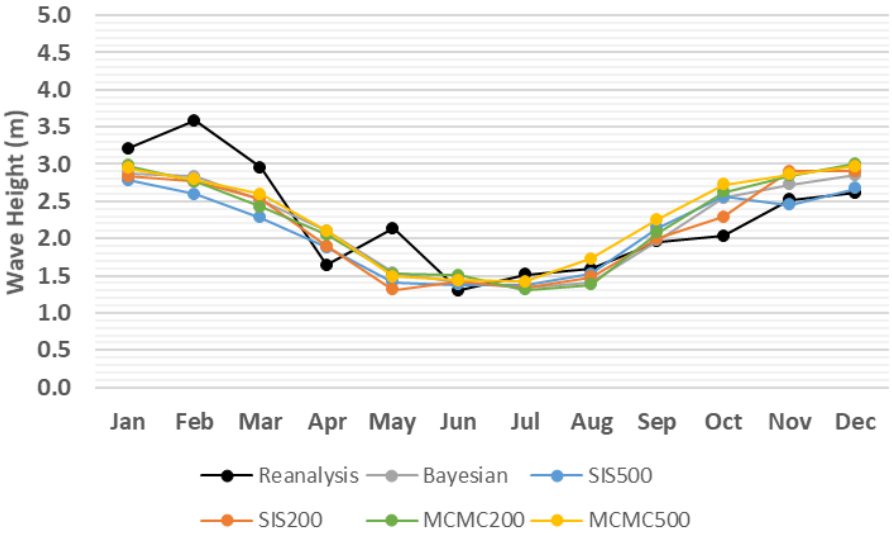


Figure 23. Monthly average comparison between simulated and reanalysis values for wave height in coordinate (74.07°N, 35.81°E) in 2012

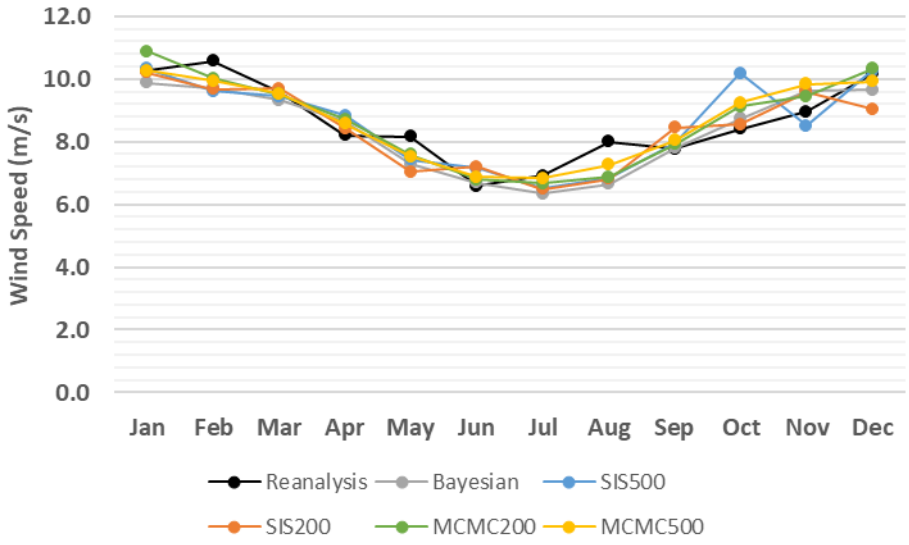


Figure 24. Monthly average comparison between simulated and reanalysis values for wind speed in coordinate (74.07°N, 35.81°E) in 2012

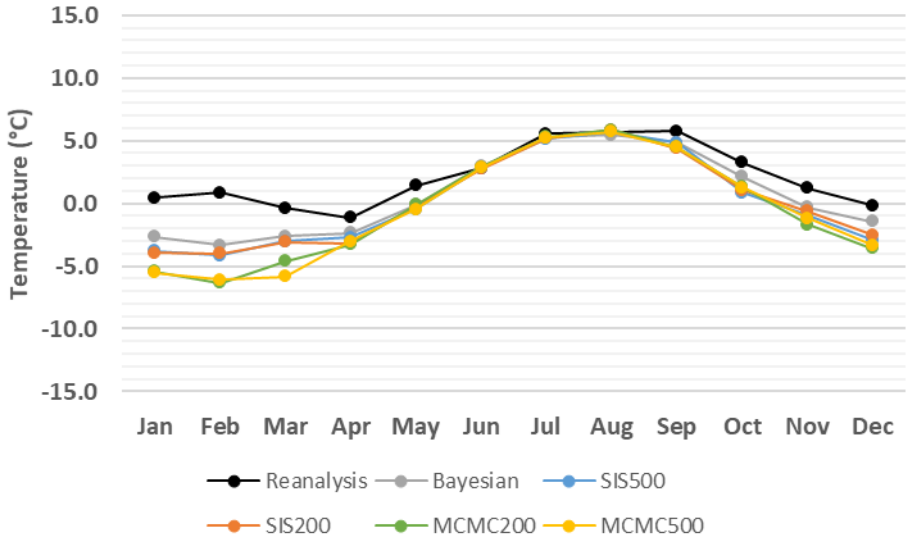


Figure 25. Monthly average comparison between simulated and reanalysis values for temperature in coordinate (74.07°N, 35.81°E) in 2012

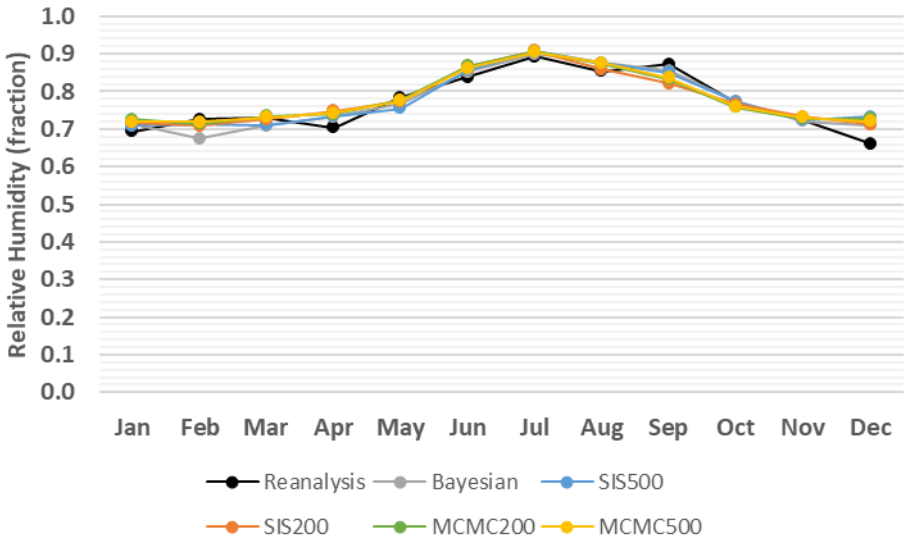


Figure 26. Monthly average comparison between simulated and reanalysis values for relative humidity in coordinate (74.07°N, 35.81°E) in 2012

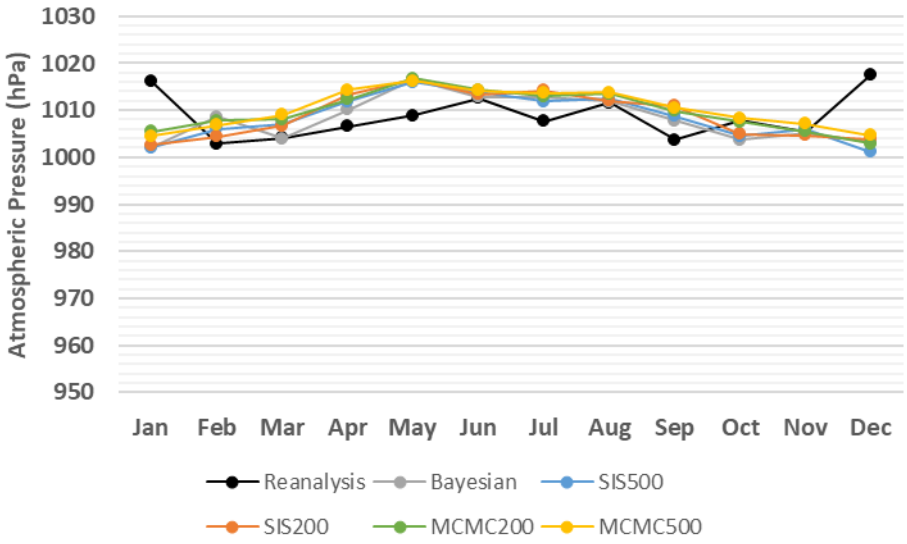


Figure 27. Monthly average comparison between simulated and reanalysis values for atmospheric pressure in coordinate (74.07°N, 35.81°E) in 2012

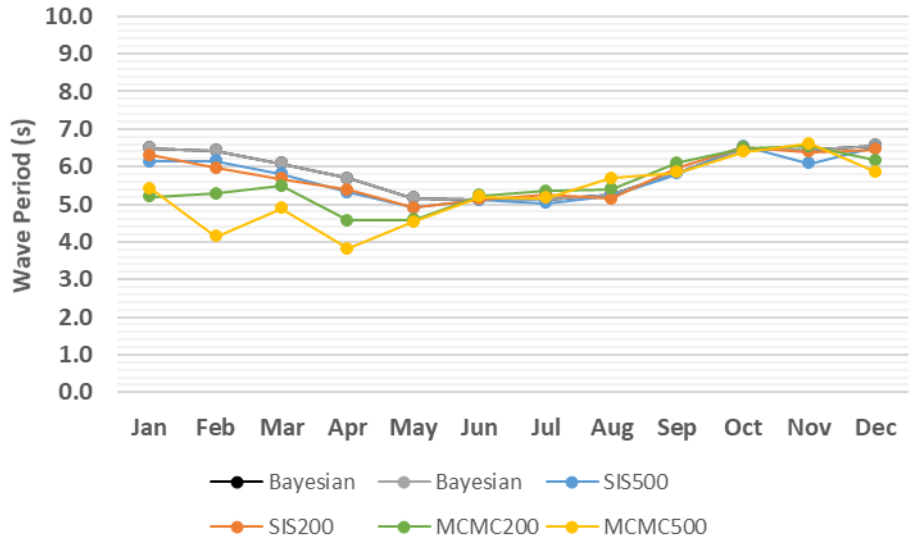


Figure 28. Monthly average comparison between simulated and reanalysis values for wave period in coordinate (74.07°N, 35.81°E) in 2012

Table 5. Monthly average of the deviation from reanalysis values for wave height (m) in coordinate (74.07°N, 35.81°E) in 2012

Month	Bayesian	SIS200	SIS500	MCMC200	MCMC500
Jan	1.00	1.03	**0.94**	0.99	0.97
Feb	**0.97**	1.19	1.25	1.00	0.97
Mar	0.89	0.96	1.12	0.97	**0.84**
Apr	0.65	0.74	**0.63**	0.67	0.70
May	0.98	1.10	1.04	0.95	**0.95**
Jun	0.54	0.62	**0.52**	0.66	0.53
Jul	0.42	**0.38**	0.46	0.47	0.52
Aug	**0.54**	0.56	0.59	0.56	0.74
Sep	**0.82**	0.83	1.02	0.89	1.07
Oct	0.99	**0.94**	1.17	1.15	1.22
Nov	**0.66**	0.84	0.75	0.78	0.83
Dec	**1.03**	1.48	1.37	1.15	1.09

Table 6. Monthly average of the deviation from reanalysis values for wind speed (m/s) in coordinate (74.07°N, 35.81°E) in 2012

Month	Bayesian	SIS200	SIS500	MCMC200	MCMC500
Jan	3.39	3.37	**3.31**	3.76	3.52
Feb	2.39	**2.23**	2.31	2.29	2.38
Mar	**2.69**	3.14	2.92	2.77	2.89
Apr	1.99	2.07	2.65	**1.92**	2.18
May	2.77	**2.50**	2.84	2.60	2.54
Jun	2.33	2.49	2.49	2.22	**2.14**
Jul	1.56	1.86	1.90	1.65	**1.53**
Aug	2.48	2.51	2.60	**2.35**	2.62
Sep	**2.90**	3.62	3.23	3.01	3.00
Oct	**2.85**	3.18	4.10	3.22	3.07
Nov	2.63	3.13	3.48	**2.43**	2.59
Dec	**2.83**	3.61	3.64	2.94	2.91

Table 7. Monthly average of the deviation from reanalysis values for temperature (°C) in coordinate (74.07°N, 35.81°E) in 2012

Month	Bayesian	SIS200	SIS500	MCMC200	MCMC500
Jan	**3.13**	4.36	4.29	5.95	5.99
Feb	**4.25**	4.88	5.16	7.25	6.94
Mar	**2.76**	3.19	3.16	4.49	5.57
Apr	**1.83**	2.58	2.26	2.40	2.18
May	1.68	1.85	2.17	**1.64**	1.98
Jun	0.63	0.64	0.76	0.67	**0.57**
Jul	0.75	0.84	**0.71**	0.80	0.76
Aug	0.85	0.83	0.97	0.89	**0.78**
Sep	1.33	1.83	**1.25**	1.43	1.42
Oct	**1.41**	2.44	2.46	2.07	2.12
Nov	**2.34**	2.75	3.02	3.38	3.04
Dec	**2.12**	3.31	3.44	3.55	3.49

Table 8. Monthly average of the deviation from reanalysis values for relative humidity (fraction) in coordinate (74.07°N, 35.81°E) in 2012

Month	Bayesian	SIS200	SIS500	MCMC200	MCMC500
Jan	5.31	5.59	**5.20**	5.82	6.02
Feb	9.54	9.34	9.33	**8.21**	8.25
Mar	5.94	7.53	7.49	6.28	**5.16**
Apr	9.72	10.39	**9.59**	9.94	9.65
May	**8.22**	8.82	8.57	8.60	8.54
Jun	5.54	5.41	5.25	4.96	**4.88**
Jul	6.48	**5.98**	7.34	6.36	6.36
Aug	6.55	6.14	7.52	**5.95**	5.99
Sep	**8.81**	10.47	9.22	9.63	9.75
Oct	6.46	8.41	9.76	**6.19**	6.21
Nov	**9.98**	13.56	11.86	11.59	11.29
Dec	**6.97**	9.38	8.16	8.97	7.67

Table 9. Monthly average of the deviation from reanalysis values for atmospheric pressure (hPa) in coordinate (74.07°N, 35.81°E) in 2012

Month	Bayesian	SIS200	SIS500	MCMC200	MCMC500
Jan	14.20	14.09	15.64	**13.07**	14.57
Feb	17.13	**15.96**	16.92	16.46	16.86
Mar	**10.58**	12.53	12.60	11.55	12.91
Apr	9.35	12.01	9.41	**8.40**	10.38
May	**9.80**	11.05	9.92	10.46	9.95
Jun	4.40	5.05	**4.20**	4.84	4.88
Jul	**7.07**	8.74	7.66	7.45	7.76
Aug	7.62	8.29	**6.44**	7.46	6.84
Sep	**8.79**	9.35	10.88	10.40	9.69
Oct	8.46	8.41	12.09	8.68	**7.70**
Nov	12.85	**12.20**	14.82	12.36	12.79
Dec	16.90	17.16	18.81	17.05	**16.12**

Table 10. Monthly average of the deviation from reanalysis values for wave period (s) in coordinate (74.07°N, 35.81°E) in 2012

Month	Bayesian	SIS200	SIS500	MCMC200	MCMC500
Jan	**1.06**	1.14	1.08	2.00	1.82
Feb	**1.14**	1.59	1.38	2.33	3.38
Mar	**0.99**	1.40	1.20	1.52	2.16
Apr	**0.78**	0.81	0.80	1.34	1.61
May	**1.03**	1.36	1.29	1.47	1.56
Jun	**0.63**	1.03	0.77	0.70	0.67
Jul	**0.70**	0.97	0.79	0.84	0.71
Aug	**0.60**	0.62	0.70	0.71	0.96
Sep	0.64	0.63	0.88	0.89	**0.61**
Oct	**0.98**	1.22	1.06	1.13	1.01
Nov	**0.63**	0.98	1.13	0.83	0.73
Dec	**0.89**	1.01	1.25	1.28	1.47

Apart from the predicted parameters, some other parameters are required in the MINCOG model that are adopted from the study by Naseri and Samuelsen (2019). Accordingly, the salinity of seawater is kept constant as 35 ppt, the ship speed is 4 m/s, and surface seawater temperature is 2.5 °C. Additionally, winds and waves are coming from the same direction and the direction between wave and the ship is 150 degree (Naseri & Samuelsen, 2019). Eventually, the predicted parameters from different algorithms are separately plugged in the MINCOG model along with the adopted parameters as inputs to estimate the daily icing rate in coordinate (74.07°N, 35.81°E) in 2012. Consequently, the monthly average of icing rate is depicted in Figure 29 in which all of the techniques have let to competitive estimates quite close to the reanalysis values. Moreover, monthly averages of deviations from reanalysis values are indicated in Table 11. Accordingly, the closest estimates to reanalysis values are related to Bayesian inputs followed by SIS200, SIS500, MCMC200, and MCMC500. Accordingly, the MINCOG estimations with Bayesian inputs do not deviate more than 0.13 cm/h from reanalysis values. While, the largest deviation, yet competitive, is 0.41 cm/h related to MCMC200 inputs in February.

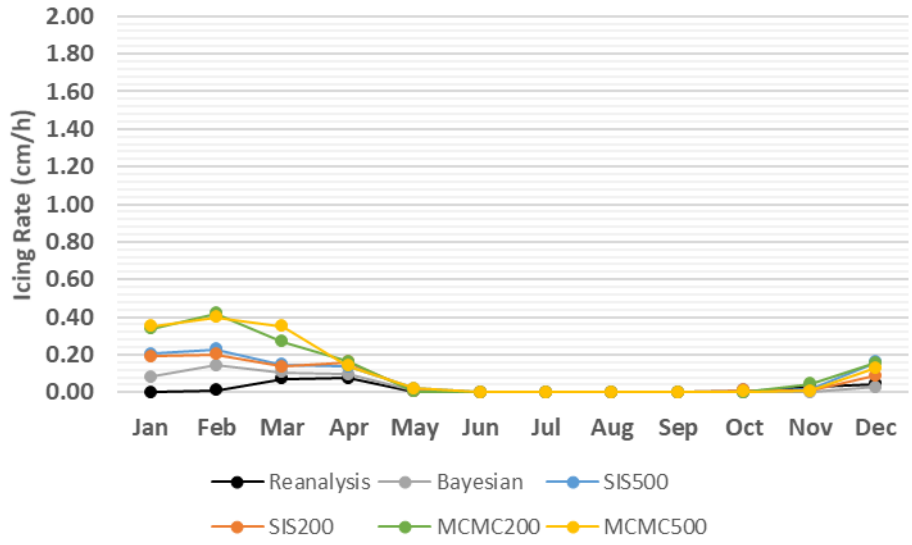


Figure 29. Monthly average of icing rate (cm/h) using MINCOG model and simulated input parameters from different algorithms comparing with reanalysis values in coordinate (74.07°N, 35.81°E) in 2012

Table 11. Monthly average of the deviation from reanalysis values for icing rate (cm/h) using MINCOG model and simulated input parameters from different algorithms in coordinate (74.07°N, 35.81°E) in 2012

Month	Bayesian	SIS200	SIS500	MCMC200	MCMC500
Jan	**0.08**	0.19	0.20	0.34	0.35
Feb	**0.13**	0.19	0.22	0.41	0.39
Mar	**0.08**	0.12	0.12	0.22	0.29
Apr	**0.07**	0.12	0.10	0.10	0.09
May	**0.00**	**0.00**	0.02	0.01	0.02
Jun	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Jul	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Aug	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Sep	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Oct	**0.00**	0.01	0.01	**0.00**	**0.00**
Nov	**0.03**	0.04	0.04	0.07	0.04
Dec	**0.06**	0.11	0.16	0.14	0.14

Additionally, the elapsed times of the Bayesian approach and simulation algorithms and the entire of the study area are illustrated in Table 12. Hence, the Bayesian framework is much faster than the other algorithms by the capability of predicting the six parameters for 366 days of the year 2012 in only 1 second for one location and 00:04:41 (i.e. 281 seconds) for the entire

area. However, the running times of the other algorithms are quite low and competitive. The worst-case in the simulation of the six parameters for one location in a year is related to SIS500 with 22 seconds. While in the simulation of the entire area MCMC500 has the longest running time with 02:22:14. It should be mentioned that reading the data from the dataset and extracting the required information is a separate task, which lasts for about 1 minute for the entire area and not affected by the algorithms. Likewise, simulating daily icing rate for one year in the entire area by the MINCOG model takes around 16 hours that clearly is not influenced by the algorithms.

Table 12. Elapsed time (hh:mm:ss) of Bayesian approach and simulation algorithms in forecasting the six meteorological and oceanographic parameters

	Bayesian	**SIS200**	**SIS500**	**MCMC200**	**MCMC500**
Coordinate (74.07°N, 35.81°E)	00:00:01	00:00:05	00:00:22	00:00:12	00:00:19
Entire area	00:04:41	00:44:29	02:00:03	01:28:28	02:22:14

4 Conclusions and Recommendations

In this study, simulation of meteorological and oceanographic parameters to improve the estimation of sea spray icing in the Arctic region was purposed. As a matter of the climate change phenomenon, direct use of old historical data may not lead to proper predictions for the future. Therefore, a Bayesian approach was considered which modifies prior belief regarding data while receiving recently sampled data. However, the prior distribution plays a key role in Bayesian inference. Therefore, the assumption that the parameters are being generated based on a Gaussian process was evaluated through the Anderson-Darling test. Eventually, the results showed that the prior distribution was reasonably modified in terms of both central tendency and deviation.

Furthermore, sampling and simulation techniques comprised of MCS, rejection sampling, IS, SIR, SIS, and MCMC were investigated. Consequently, the standard SIS and MCMC were modified and the proposed algorithms were developed. Then, using two sample sizes of 200 and 500, the performance of the proposed SIS and MCMC were examined. In the proposed MCMC; however, another stopping criterion was also considered to evaluate the variation of the new draws. Meanwhile, 32 years (1980-2011) of reanalysis data from NORA10 was used

to estimate the kernel smoothing density as the target density. Moreover, Weibull distribution was considered as the proposal density. However, since Weibull distribution is only defined for positive values, a data shifting procedure was embedded in the algorithms. Then, four combinations of the proposed algorithms and sample sizes (i.e. SIS200, SIS500, MCMC200, and MCMC500), as well as the Bayesian framework, six meteorological and oceanographic parameters, were simulated for the year 2012 on a daily basis. Accordingly, all the algorithms reached competitive results while the Bayesian model indicated slightly lower deviations from the reanalysis values in 2012. The Bayesian model was also much faster with the capability of predicting the six parameters for the entire area in less than five minutes. Therefore, the Bayesian approach is considered to be the most resistant technique, which is also robust due to few required assumptions for implementation. Moreover, the results implied that the proposed SIS and MCMC could properly cope with simulating the parameters in quite low number of iterations (i.e. 200) while further iterations result in only little improvements. Hence, although SIS200 and MCMC200 take a longer time rather than Bayesian framework, their running times for the entire area (i.e. less than 1 hour and 1.5 hours, respectively) are yet reasonable.

Eventually, the simulated values from all algorithms were considered as inputs in the MINCOG model to forecast the daily icing rate in 2012. Accordingly, the best results were obtained using Bayesian inputs closely followed by SIS200, SIS500, MCMC200, and MCMC500.

The applied approaches and proposed models can play useful roles in industrial application, especially, when new data and information are collected using which the meteorological and atmospheric conditions are predicted for future junctures. This provides the decision-maker with valuable information for planning offshore activities in the future (e.g., offshore fleet optimization). Sea voyages with relatively lower risks can be selected based on the predicted icing rates. Further works can focus on developing a Bayesian approach using an empirical prior distribution rather than a Gaussian data-generating process. Besides, the models can be combined with Ant Colony Optimization (ACO), which is a promising approach to routing problems, in a Multi-Criteria Decision-Making (MCDM) framework to determine the shortest and safest sea routes for vessels. Moreover, the models can be embedded in planning and scheduling problems, such as maintenance scheduling in a dynamic condition to predict the possible intervals for maintenance activities.

Works Cited

Ackoff, R. (1989). From Data to Wisdom. *Journal of Applied Systems Analysis, 16*, 8-9.

Ali, M., Deo, R. C., Downs, N. J., & Maraseni, T. (2020). Chapter 3 - Monthly rainfall forecasting with Markov Chain Monte Carlo simulations integrated with statistical bivariate copulas. In P. Samui, D. T. Bui, S. Chakraborty, & R. C. Deo, *Handbook of Probabilistic Models* (pp. 89-105). Oxford, UK: Elsevier Inc. doi:10.1016/B978-0-12-816514-0.00003-5

Anderson, T. W., & Darling, D. A. (1952). Asymptotic theory of certain 'goodness-of-fit' criteria based on stochastic processes. *The Annals of Mathematical Statistics, 23*(2), 193-212.

Barabadi, A., Garmabaki, A., & Zaki, R. (2016). Designing for performability: An icing risk index for Arctic offshore. *Cold Regions Science and Technology, 124*, 77-86. doi:10.1016/j.coldregions.2015.12.013

Brooks, S. P., & Roberts, G. O. (1998). Convergence assessment techniques for Markov chain Monte Carlo. *Statistics and Computing, 8*, 319±335. doi:10.1023/A:1008820505350

Chatterton, M., & Cook, J. C. (2008). *The Effects of Icing on Commercial Fishing Vessels.* Worcester, UK: Worcester Polytechnic Institute. Retrieved May 2, 2020

Cornejo-Bueno, L., Garrido-Merchán, E., Hernández-Lobato, D., & Salcedo-Sanz, S. (2018, January 31). Bayesian optimization of a hybrid system for robust ocean wave features prediction. *Neurocomputing, 275*, 818-828. doi:10.1016/j.neucom.2017.09.025

Cowles, M. K., & Carlin, B. P. (1996). Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review. *Journal of the American Statistical Association, 91*(434), 883-904. doi:10.2307/2291683

Dehghani-Sanij, A. R., Dehghani, S. R., Naterer, G. F., & Muzychka, Y. S. (2017, March 1). Sea spray icing phenomena on marine vessels and offshore structures: Review and formulation. *Ocean Engineering, 132*, 25-39. doi:10.1016/j.oceaneng.2017.01.016

Dehghani-sanij, A., Muzychka, Y. S., & Naterer, G. F. (2015). Analysis of Ice Accretion on Vertical Surfaces of Marine Vessels and Structures in Arctic Conditions. *Proceedings of the ASME 2016 35th International Conference on Ocean, Offshore and Arctic Engineering. 7.* St. John's, Newfoundland, Canada: ASME. doi:10.1115/OMAE2015-41306

Elsner, J. B., & Bossak, B. H. (2001, December 1). Bayesian Analysis of U.S. Hurricane Climate. *Journal of Climate, 14*, 4341-4350. Retrieved October 25, 2019, from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.490.4967&rep=rep1&type=pdf

Epstein, E. S. (1985). Statistical Inference and Prediction in Climatology: A Bayesian Approach. *Meteorological Monograph, 20*(42), 199.

Flickr. (2019, February 1). *Flickr*. Retrieved December 10, 2019, from Flickr: https://www.flickr.com/photos/7455207@N05/47106204561/in/photostream/

Forest, T., Lozowski, E., & Gagnon, R. E. (2005). Estimating Marine Icing on Offshore Structures Using RIGICE04. *the 11th International Workshop on Atmospheric Icing of Structures* (pp. 12-16). Montreal, QC, Canada: National Research Council Canada. Retrieved May 4, 2020

Gilks, W. R., & Berzuini, C. (2001). Following a moving target—Monte Carlo inference for dynamic Bayesian models. *Journal of the Royal Statistical Society B, 63*(1), 127-146. doi:10.1111/1467-9868.00280

Givens, G. H., & Hoeting, J. A. (2013). *Computational Statistics* (2nd ed.). John Wiley & Sons Inc. doi:10.1002/9781118555552

Harter, H. L. (1984). Another look at plotting positions. *Communications in Statistics - Theory and Methods, 13*(13), 1613-1633. doi:10.1080/03610928408828781

Heinrich, H. (1950). *Industrial Accident Prevention: A Scientific Approach.* (3rd, Ed.) McGraw Hill.

Horjen, I. (1960). *Numerical Modelling of Time-dependent Marine Icing, Anti-icing and De-icing.* Trondheim, Norway: Norges Tekniske Høgskole (NTH).

Horjen, I. (2013, September). Numerical modeling of two-dimensional sea spray icing on vessel-mounted cylinders. *Cold Regions Science and Technology, 93*, 20-35. doi:10.1016/j.coldregions.2013.05.003

Kulyakhtin, A., & Tsarau, A. (2014). A time-dependent model of marine icing with application of computational fluid dynamics. *Cold Regions Science and Technology, 104-105*, 33-44. doi:10.1016/j.coldregions.2014.05.001

Lee, P. M. (1997). *Bayesian Statistics, an Introduction* (2nd ed.). Wiley.

Little, R. J. (2006). Calibrated Bayes: A Bayes/Frequentist Roadmap. *The American Statistician, 60*(3), 213-223. doi:10.1198/000313006X117837

Liu, J. S., & Chen, R. (1988). Sequential Monte Carlo Methods for Dynamic Systems. *Journal of the American Statistical Association, 93*(443), 1032-1044. doi:10.2307/2669847

Lozowski, E. P. (n.d.). Sea Spray Icing Of Ships And Offshore Structures. In *Cold region science and marine technology.* Encyclopedia of Life Support Systems (EOLSS). Retrieved December 10, 2019, from https://www.eolss.net/Sample-Chapters/C05/E6-178-62.pdf

Lozowski, E., Szilder, K., & Lasse Makkonen, L. (2000, November 15). Computer simulation of marine ice accretion. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 358*, 2811–2845. doi:10.1098/rsta.2000.0687

Marine Insight . (2020). *Safety of Life at Sea (SOLAS) – The Ultimate Guide*. Retrieved May 13, 2020, from https://www.marineinsight.com/maritime-law/safety-of-life-at-sea-solas-convention-for-prevention-of-marine-pollution-marpol-a-general-overview/

Mertins, H. O. (1968). Icing on fishing vessels due to spray. *Marine Observer, 38*(221), 128-130. Retrieved December 10, 2019

Naseri, M., & Samuelsen, E. M. (2019). Unprecedented Vessel-Icing Climatology Based on Spray-Icing Modelling and Reanalysis Data: A Risk-Based Decision-Making Input for Arctic Offshore Industries. *Atmosphere, 10*(4), 197. doi:10.3390/atmos10040197

Naseri, M., Baraldi, P., Compare, M., & Zio, E. (2016). Availability assessment of oil and gas processing plants operating under dynamic Arctic weather conditions. *Reliability Engineering and System Safety, 152*, 66–82. doi:10.1016/j.ress.2016.03.004

Nikam, V. B., & Meshram, B. (2013). Modeling Rainfall Prediction Using Data Mining Method: A Bayesian Approach. *Fifth International Conference on Computational Intelligence, Modelling and Simulation* (pp. 132-136). Seoul: IEEE. doi:10.1109/CIMSim.2013.29

Norwegian Petroleum Directorate. (2020). *FactPages*. (Norwegian Petroleum Directorate) Retrieved May 4, 2020, from https://factpages.npd.no/en

Overland, J. (2000, June 21). *Naval Postgraduate School*. Retrieved December 10, 2019, from https://www.met.nps.edu/~psguest/polarmet/vessel/ves_photos.html

Park, M. H., Ju, M., & Kim, J. Y. (2020). Bayesian approach in estimating flood waste generation: A case study in South Korea. *Journal of Environmental Management, 265*, 110552. doi:10.1016/j.jenvman.2020.110552

Pole, A., West, M., & Harrison, J. (1994). *Applied Bayesian Forecasting and Time Series Analysis* (1st ed.). New York: Chapman and Hall/CRC. doi:10.1201/9781315274775

Rajagopalan, B., Lall, U., & Tarboton, D. G. (1997). Evaluation of kernel density estimation methods for daily precipitation resampling. *Stochastic Hydrology and Hydraulics*, *11*, 523-547. doi:10.1007/BF02428432

Rashid, T., Khawaja, H. A., & Edvardsen, K. (2016). Review of marine icing and anti-/de-icing systems. *Journal of Marine Engineering & Technology, 15*(2), 79-87. doi:10.1080/20464177.2016.1216734

Reistad, M., Breivik, Ø., Haakenstad, H., Aarnes , O. J., Furevik, B., & Bidlot, J.-R. (2011, May 26). A high-resolution hindcast of wind and waves for the North Sea, the Norwegian Sea, and the Barents Sea. *Journal of Gephysical Research: Ocean, 116*(C05019). doi:10.1029/2010JC006402

Ridgeway, G., & Madigan, D. (2003). A Sequential Monte Carlo Method for Bayesian Analysis of Massive Datasets. *Data Mining and Knowledge Discovery, 7*, 301–319.

Robert, C., & Casella, G. (2012, January 9). A Short History of Markov Chain Monte Carlo: Subjective Recollections from Incomplete Data. *Statistical Science, 26*(1), 102–115. doi:10.1214/10-STS351

Rubin, D. B. (1981). The Bayesian bootstrap. *Annals of Statistics, 9*(1), 130-134. doi:10.1214/aos/1176345338

Rubin, D. B. (1987). A Noniterative Sampling/Importance Resampling Alternative to the Data Augmentation Algorithm for Creating a Few Imputations When Fractions of Missing Information Are Modest: The SIR Algorithm. *Journal of the American Statistical Association, 82*(398), 543-546. doi:10.1080/01621459.1987.10478461

Rubin, D. B. (1988). Using the SIR algorithm to simulate posterior distributions. In J. M. Bernardo, M. H. DeGroot, D. V. Lindley, & A. F. Smith, *Bayesian Statistics 3* (pp. 395-402). Oxford: John Wiley & Sons, Inc.

Ryerson, C. C. (2011, January). Ice protection of offshore platforms. *Cold Regions Science and Technology, 65*(1), 97–110. doi:10.1016/j.coldregions.2010.02.006

Samuelsen, E. M. (2017, October 11). Ship-icing prediction methods applied in operational weather forecasting. *Quarterly Journal of the Royal Meteorological Society, 144*(710), 13-33. doi:10.1002/qj.3174

Samuelsen, E. M. (2018). Ship-icing prediction methods applied in operational weather forecasting. *Quarterly Journal of the Royal Meteorological Society, 144*, 13-33. doi:10.1002/qj.3174

Samuelsen, E. M., & Graversen, R. G. (2019). Weather situation during observed ship-icing events off the coast of Northern Norway and the Svalbard archipelago. *Weather and Climate Extremes, 24*, 100200. doi:10.1016/j.wace.2019.100200

Samuelsen, E. M., Edvardsen, K., & Graversen, R. G. (2017). Modelled and observed sea-spray icing in Arctic-Norwegian waters. *Cold Regions Science and Technology, 134*, 54-81. doi:10.1016/j.coldregions.2016.11.002

SARex1. (2016). *Search and rescue exercise conducted off north Spitzbergen.* (K. E. Solberg, O. T. Gudmestad, & B. O. Kvamme, Eds.) Stavanger: University of Stavanger. Retrieved from http://hdl.handle.net/11250/2414815

SARex2. (2017). *Surviving a maritime incident in cold climate conditions.* (K. E. Solberg, O. T. Gudmestad, & E. Skjærseth, Eds.) Stavanger: University of Stavanger. Retrieved from https://brage.bibsys.no/xmlui/handle/11250/2468805

SARex3. (2018). *Evacuation to shore, survival and rescue.* (K. E. Solberg, & O. T. Gudmestad, Eds.) Stavanger: University of Stavanger. Retrieved from https://brage.bibsys.no/xmlui/handle/11250/2578301

Scott, D. W. (2015). *Multivariate Density Estimation: Theory, Practice, and Visualization* (2nd ed.). New York: Wiley.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis.* London: Chapman & Hall.

Stallabrass, J. R. (1980). *Trawler Icing: A Compilation of Work Done at N.R.C.* Ottawa: National Research Conseil Canada. Retrieved December 10, 2019, from https://apps.dtic.mil/dtic/tr/fulltext/u2/a100832.pdf

Sultana, K., Dehghani, S., Pope, K., & Muzychka, Y. (2018). A review of numerical modelling techniques for marine icing applications. *Cold Regions Science and Technology, 145*, 40-51. doi:10.1016/j.coldregions.2017.08.007

Tierney, L. (1994). Markov Chains for Exploring Posterior Distributions. *Annals of Statistics, 22*(4), 1701-1762. doi:10.1214/aos/1176325750

Toomey, R. M., Lloyd, M., House, D. J., & Dickins, D. (2010). *The Ice Navigation Manual.* Edinburgh, UK: Witherby Seamanship International Ltd.

Tukey, J. W. (1977). *Exploratory Data Analysis.* Reading, Massachusetts: Addison-Wesley.

Walshaw, D. (2000). Modelling extreme wind speeds in regions prone to hurricanes. *Applied Statistics, 49*(1), 51-62. doi:10.1111/1467-9876.00178

Wang, J., Fonseca, R. M., Rutledge, K., Martín-Torres, J., & Yu, J. (2019, March). Weather Simulation Uncertainty Estimation Using Bayesian Hierarchical Models. *Journal of Applied Meteorology and Climatology, 58*, 585-603. doi:10.1175/JAMC-D-18-0018.1

Wikle, C. K., Milliff, R. F., Herbei, R., & Leeds, W. B. (2013). Modern Statistical Methods in Oceanography: A Hierarchical Perspective. *Statistical Science, 28*(4), 466–486. doi:10.1214/13-STS436

Wilks, D. (2011). *Statistical Methods in the Atmospheric Sciences* (3rd ed., Vol. 100). Academic Press. Retrieved October 25, 2019, from https://www.elsevier.com/books/statistical-methods-in-the-atmospheric-sciences/wilks/978-0-12-385022-5

Zhao, N., Lu, N., Chen, C., Li, H., Yue, T., Zhang, L., & Liu, Y. (2017). Mapping temperature using a Bayesian statistical method and a high accuracy surface modelling method in the Beijing–Tianjin–Hebei region, China. *Royal Meteorological Society, 24*(4), 571-579. doi:10.1002/met.1657

Zio, E. (2013). *The Monte Carlo Simulation Method for System Reliability and Risk Analysis.* London: Springer. doi:10.1007/978-1-4471-4588-2